PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 :
C12N 15/12, C07K 14/705, 16/28, A61K 38/17 — A2

(11) International Publication Number: WO 00/18915

(43) International Publication Date: 6 April 2000 (06.04.00)

(21) International Application Number: PCT/US99/22082

(22) International Filing Date: 23 September 1999 (23.09.99)

(30) Priority Data:

60/155,215	25 September 1998 (25.09.98)	US
60/155,251	13 October 1998 (13.10.98)	US
60/172,228	4 May 1999 (04.05.99)	US

(63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Applications

US	60/155,251 (CIP)
Filed on	13 October 1998 (13.10.98)
US	60/172,228 (CIP)
Filed on	4 May 1999 (04.05.99)
US	60/155,215 (CIP)
Filed on	25 September 1998 (25.09.98)

(71) Applicant *(for all designated States except US)*: INCYTE PHARMACEUTICALS, INC. [US/US]; 3174 Porter Drive, Palo Alto, CA 94304 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: YUE, Henry [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US). LAL, Preeti [IN/US]; 2382 Lass Drive, Santa Clara, CA 9054 (US). CORLEY, Neil, C. [US/US]; 1240 Dale Avenue #30, Mountain View, CA 94040 (US). GUEGLER, Karl, J. [CH/US]; 1048 Oakland Avenue, Menlo Park, CA 94025 (US). BAUGHN, Mariah, R. [US/US]; 14244 Santiago Road, San Leandro, CA 94577 (US). LU, Aina, D. [US/US]; 55 Park Belmont Place, San Jose, CA 95136 (US). TANG, Y., Tom [CN/US]; 4230 Ranwick Court, San Jose, CA 95118 (US).

(74) Agents: BILLINGS, Lucy, J. et al.; Incyte Pharmaceuticals, Inc., 3174 Porter Drive, Palo Alto, CA 94304 (US).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: MEMBRANE-ASSOCIATED ORGANIZATIONAL PROTEINS

```
1   M S T A R E Q P I F S T R A H V F Q I D  2687924
1   M G K M G E Q P I F S T R A H V F Q I D  g1913909

21  P A T K R N W I P A G K H A L T V S Y F  2687924
21  P N T K K N W V P T S K H A V T V S Y F  g1913909

41  Y D A T R N V Y R I I S I G G A K A I I  2687924
41  Y D S T R N V Y R I I S L D G S K A I I  g1913909

61  N S T V T P N M T F T K T S Q K F G Q W  2687924
61  N S T I T P N M T F T K T S Q K F G Q W  g1913909

81  A D S R A N T V Y G L G F A S E Q H L T  2687924
81  A D S R A N T V Y G L G F S S E H H L S  g1913909

101 Q F A E K F Q E V K E A A R L A R E K S  2687924
101 K F A E K F Q E F K E A A R L A K E K S  g1913909

121 Q D G G E L T S P A L G L A S H Q V P P  2687924
121 Q E K M E L T S T - - - - - - - - - - -  g1913909
```

(57) Abstract

The invention provides human membrane-associated organizational proteins (HJNCT) and polynucleotides which identify and encode HJNCT. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating, or preventing disorders associated with expression of HJNCT.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

MEMBRANE-ASSOCIATED ORGANIZATIONAL PROTEINS

TECHNICAL FIELD

This invention relates to nucleic acid and amino acid sequences of human membrane-associated organizational proteins and to the use of these sequences in the diagnosis, treatment, and prevention of cell proliferative disorders, including cancer, and autoimmune/inflammatory, neurological, developmental, vesicle trafficking, reproductive, gastrointestinal, and renal disorders.

BACKGROUND OF THE INVENTION

Cells are surrounded by plasma membranes which enclose the cell and maintain an environment inside the cell that is distinct from its surroundings. Eukaryotic organisms are distinct from prokaryotes in that they possess many intracellular organelle and vesicle structures enclosed by membranes. Membrane-associated organizational proteins are responsible for the aggregation and assembly of signaling and transport proteins at specialized regions of cellular, organelle, and vesicular membranes. For example, in postsynaptic signaling, membrane-associated organizational proteins are responsible for ion channel and receptor clustering. In Golgi-mediated transport and secretion, membrane-associated organizational proteins control cisternal stacking and vesicle docking. Membrane-associated organizational proteins also play a role in the formation of cell junctions, regions of contact between adjacent cells and between cells and the extracellular matrix. Cell junctions influence cell shape, strength, flexibility, motility, and adhesion.

PDZ Domains

A conserved protein domain called PDZ has been identified in various proteins which act at the cytosolic face of the plasma membrane. PDZ-containing proteins coordinate the assembly of multifunctional protein complexes involved in intercellular signaling events. PDZ domains are protein/protein interaction motifs involved, for example, in the localization of channels, receptors, signaling enzymes, and adhesion molecules to sites of cell-cell contact. PDZ domains were named for three proteins in which this domain was initially discovered: PSD-95 (postsynaptic density 95), Dlg (Drosophila lethal(1)discs large-1), and ZO-1 (zonula occludens-1). These proteins play important roles in neuronal synaptic transmission, tumor suppression, and cell junction formation, respectively. Since the discovery of these proteins, over sixty additional PDZ-containing proteins have been identified in diverse prokaryotic and eukaryotic organisms. A large proportion of PDZ domains are found in the eukaryotic MAGUK (membrane-associated guanylate kinase) protein family, members of which bind to the intracellular domains of neuronal receptors and channels. However, PDZ domains are also found in

diverse membrane-localized proteins such as protein tyrosine phosphatases, serine/threonine kinases, G-protein cofactors, and synapse-associated proteins such as syntrophins and neuronal nitric oxide synthase (nNOS). Generally, about one to three PDZ domains are found in a given protein, although up to nine have been identified in a single protein. (See, e.g., Ponting, C.P. et al. (1997) Bioessays 19:469-479; Fanning, A.S. and J.M. Anderson (1999) J. Clin. Invest. 103:767-772.)

X-ray crystallography has shown that PDZ domains are generally compact globular structures containing about 80 to 100 amino acids which form six β-strands and two α-helices. PDZ domains tend to be rich in glycine residues which introduce turns in the polypeptide chain and promote compaction and stability of the folded polypeptide. A GLGF (glycine-leucine-glycine-phenylalanine) sequence motif is conserved within some PDZ domains. This GLGF sequence is usually preceded by an arginine found about six residues upstream of GLGF. PDZ domains bind to a tripeptide motif containing valine and serine or threonine. Most ligands which bind PDZ domains contain this motif, although some ligands lack this motif or contain conservative substitutions therein.

PDZ-containing proteins are likely involved in disorders associated with defective cell signaling (Ponting, supra). For example, PDZ domains have been shown to play important roles in development, and the gene encoding the PDZ-containing protein LIM kinase 1 is deleted in patients with Williams syndrome, a complex developmental disorder.

PDZ-Mediated Neuronal Signaling

Cells communicate with and respond to their environment by receiving and processing extracellular signals. These signals take the form of growth factors, hormones, cytokines, and peptides which bind to activate specific plasma membrane receptors. The activated receptors trigger intracellular signal transduction pathways which culminate in a wide range of cellular responses affecting gene expression, protein secretion, cell cycle progression, and cell differentiation. Initial events in signal transduction require the proximity of intracellular signaling proteins to the cytosolic domains of activated plasma membrane receptors. These intracellular membrane-associated signaling proteins couple the activated receptor to downstream second messenger systems and play a key role in the regulation and coordination of complex, multiprotein signal transduction pathways.

PDZ proteins play an important role in the clustering of ion channels and neurotransmitter receptors at postsynaptic membranes. This organizational activity is essential for neuronal development and synaptic plasticity (Ponting, supra). Mutations that block clustering of neuronal receptors and channels cause perinatal lethality in mice. MAGUK proteins, in particular, are important for clustering neuronal receptors and ion channels responsive to glutamate, the predominant excitatory neurotransmitter in the mammalian hippocampus. Specifically, the PDZ domains of PSD-95, PSD-93, SAP-97 (synapse-

associated protein 97), SAP-102, and chapsyn 110 bind to the cytosolic C-termini of N-methyl-D-aspartate (NMDA) glutamate receptors and Shaker-type potassium channels, causing them to cluster.

A novel synaptic PDZ protein, Homer, has recently been identified in rat brain. Although Homer may perform similar functions as MAGUK proteins, it is highly divergent from MAGUK proteins and may represent a novel and distinct PDZ protein family. Homer mRNA is 6.5 kb long and encodes a protein of 186 amino acids. Homer contains a single PDZ-like domain and binds to the carboxy terminus of phosphoinositide-linked metabotropic glutamate receptors. The PDZ-like domain contains a GLGF sequence and preceding arginine, as seen in the PDZ domains of proteins such as PSD-95. Otherwise, there is less than 10% amino acid sequence identity between Homer and the reported members of the PDZ family. Deletion constructs revealed that the amino-terminal 108 amino acids of Homer, which includes the GLGF sequence, is essential for the binding of Homer to glutamate receptors. Expression of Homer mRNA is strongly upregulated in the forebrain by seizure-and drug induced neuronal activation (Brakeman, P.R. et al. (1997) Nature 386:284-288).

Detailed immunohistochemical analysis revealed that Homer is enriched at excitatory synapses. Additionally, expression of the Homer gene is developmentally regulated by synaptic activity, with peak expression in the rat forebrain coinciding with increased synaptic activity from the third to fifth postnatal weeks. In the adult, Homer mRNA is rapidly induced in the hippocampus of awakened rats. Homer may also be linked to the regulation of dopamine receptors since it is rapidly induced by cocaine in the striatum (Brakeman, supra).

Taken together, these observations suggest that Homer may play a major role in neuronal function and development. Homer is likely to participate in signal transduction and influence spatial targeting of receptors. The selective expression of Homer at excitatory synapses strongly supports a role in synapse formation and in the regulation of glutamate mediated neurotransmission.

Cisternal Stacking in Golgi

The Golgi apparatus (Golgi), an organelle composed of stacked disc-shaped cisternal membranes, is found adjacent to the nucleus during interphase in animal cells. The Golgi contains enzymes that modify secreted and membrane proteins posttranslationally as they traverse the secretory pathway. Many of the modifying enzymes function in an ordered sequence, and have unique distributions within the Golgi, suggesting that Golgi structure is important for their function. Secreted and membrane proteins are transported through the Golgi in specialized vesicles which bud from the donor membrane and then dock with and fuse to the target membrane. Cisternal stacking may be viewed as a specialized form of docking event, in which one cisterna docks with another without fusion. During mitosis, the Golgi breaks down into many small vesicles and tubules that are partitioned to the two daughter cells. Vesicle fusion is

inhibited during mitosis, possibly due to an inhibition of vesicle docking. As the cell exits mitosis, the Golgi cisternae reform and restack, and vesicle transport resumes.

Vesicles must be tethered to the target membrane prior to docking and fusion. Giantin, a protein on the vesicle surface, binds to p115, a cytosolic protein that is in turn bound to its receptor GM130 on the cytosolic surface of the Golgi. GM130 is a tightly associated peripheral membrane protein which has a coiled-coil, elongated rod-like structure and appears to be a dimer. During mitosis, GM130 is phosphorylated by a cyclin-dependent kinase. This phosphorylation inhibits the binding of p115 to GM130, leading to a block in vesicle tethering and docking, and thus a block in vesicle fusion.

Golgi ReAssembly Stacking Protein (GRASP65) was identified as a Golgi protein which may play roles in the initiation and/or maintenance of cisternal stacking and in vesicle docking. Antibodies to GRASP65 prevent cisternal stacking in a cell-free system. The rat GRASP65 gene, which is expressed in all rat tissues tested, encodes a 65 kDa peripheral membrane protein which appears to exist as a dimer. Homologs of GRASP65 exist in the yeasts Saccharomyces cerevisiae and Schizosaccharomyces pombe.

GRASP65 interacts with GM130 under both mitotic and nonmitotic conditions. The GRASP65/GM130 complex can interact with p115, suggesting that all three proteins function as a complex with roles in vesicle docking and cisternal stacking. Like GM130, GRASP65 is heavily phosphorylated during mitosis. GRASP65 is myristoylated at its N-terminus, which may account for its association with the Golgi membrane. The protein contains two imperfectly repeated domains (from F16 through S108 and from W112 through P202 in GI 4432587; SEQ ID NO:10) followed by a serine-rich C-terminal domain which may be involved in cell cycle regulation. The binding site for GM130 has been mapped to amino acids G194 through I201 within the second repeated domain. The first repeated domain, though similar to the second, does not bind GM130. The binding site for GRASP65 on GM130 was mapped to the C-terminus of GM130, including the final four hydrophobic amino acids. The interaction of GRASP65 with GM130 resembles the binding of PDZ domain-containing proteins to their ligands. The GYGY sequence within the GRASP65 binding site for GM130 is similar to the conserved GLGF sequence of PDZ domains. In addition, PDZ proteins recognize the C-terminal four amino acids of their ligands. (See, e.g., Warren, G. and V. Malhotra (1998) Curr. Opin. Cell Biol. 10:493-498; Lowe, M. et al. (1998) Trends Cell Biol. 8:40-44; Barr, F.A. et al. (1997) Cell 91:253-262; Barr, F.A. et al. (1998) EMBO J. 17:3258-3268.)

Defects in protein trafficking to organelles or the cell surface are involved in numerous human diseases and disorders. Defects in the trafficking of membrane-bound receptors and ion channels are associated with cystic fibrosis (cystic fibrosis transmembrane conductance regulator), glucose-galactose malabsorption syndrome (Na^+/glucose cotransporter), hypercholesterolemia (low-density lipoprotein

receptor), and forms of diabetes mellitus (insulin receptor). Abnormal hormonal secretion is linked to disorders including diabetes insipidus (vasopressin), hyper- and hypoglycemia (insulin, glucagon), Grave's disease and goiter (thyroid hormone), and Cushing's and Addison's diseases (adrenocorticotropic hormone (ACTH)).

Cancer cells secrete excessive amounts of hormones or other biologically active peptides. Disorders related to this excessive secretion include: fasting hypoglycemia due to increased insulin secretion from insulinoma-islet cell tumors; hypertension due to increased epinephrine and norepinephrine secreted from pheochromocytomas of the adrenal medulla and sympathetic paraganglia; and carcinoid syndrome, which includes abdominal cramps, diarrhea, and valvular heart disease, caused by excessive amounts of vasoactive substances (serotonin, bradykinin, histamine, prostaglandins, and polypeptide hormones) secreted from intestinal tumors. Ectopic synthesis and secretion of biologically active peptides occurs including ACTH and vasopressin in lung and pancreatic cancers, parathyroid hormone in lung and bladder cancers, calcitonin in lung and breast cancers, and thyroid-stimulating hormone in medullary thyroid carcinoma.

Cell Junctions

Cell junctions are regions of contact between adjacent cells and between cells and the surrounding extracellular matrix. Cell junctions are comprised of both intracellular and extracellular protein complexes associated with the plasma membrane. In addition, cytoskeletal filaments that traverse the cytoplasm are anchored to the cell cortex by means of human cell junction proteins. Cell junctions are dynamic structures that are responsive to signals such as cytokines and growth factors and are also capable of signal transduction. The dynamic properties of cell junctions influence cell shape, strength, flexibility, motility, and adhesion. Cell-cell and cell-matrix contacts are often disrupted in neoplastically transformed cells, suggesting a mechanism for uncontrolled cell proliferation and metastasis.

Tight junctions are present around the lateral circumference of epithelial or endothelial cells. (Reviewed in Balda, M.S. and K. Matter (1998) J. Cell Sci. 111:541-547; Lampugnani, M.G. and E. Dejana (1997) Curr. Opin. Cell Biol. 9:674-682.) Epithelia and endothelia are monolayers of polarized cells that separate a body compartment (the basolateral side) from the outside environment or a topologically equivalent space (the apical side). The apical and basolateral domains are polarized, containing different cell membrane components such as lipids and membrane-associated proteins. Tight junctions constitute a continuous, circumferential seal around cells, forming a barrier to diffusion of solutes across the cell sheet. Tight junctions also function as a boundary between apical and basolateral membrane domains, preventing lateral diffusion of membrane associated proteins, such as receptors, between compartments and thus maintaining cell polarity.

Maintenance of epithelial cell polarity is essential for the proper function of many epithelial organs. In the kidney, for example, the functions of reabsorption and secretion depend upon the polarized insertion of specialized channels and transporters to apical membranes lining the renal tubule lumen or basolateral membranes adjacent to the insterstitium and blood space. Defective polarization of membrane proteins can lead to renal cystic diseases. (Wilson, P.D. (1997) Am. J. Physiol. 272:F434-F442.) The barrier function of tight junctions is also important, as disruptions to tight junction permeability are involved in a wide range of gastrointestinal pathologies. Agents such as aspirin or ethanol which increase gastric tight junction permeability initiate and amplify gastric mucosal injury by allowing back-diffusion of H^+ ions into the mucosa. Abnormal tight junction permeability may also be the cause of inflammatory bowel disease such as Crohn's disease. Certain bacterial toxins cause intestinal epithelial tight junction abnormalities and contribute to diarrhea by dissipating electrochemical gradients needed for proper intestinal absorbtion and secretion. Structural and functional disruptions of tight junctions are also observed in inherited cholestatic liver disorders and cholestasis associated with common bile duct obstruction of the liver (Balda, M.S. et al. (1992) Yale J. Biol. Med. 65:725-735).

The protein components of tight junctions include ZO-1 and ZO-2 (zona occludens), cytoplasmic proteins associated with the plasma membrane at tight junctions. ZO-1 is a PDZ domain-containing protein which associates with spectrin and thus may link tight junctions to the actin cytoskeleton. Other cytoplasmic components of tight junctions include cingulin, 7H6 antigen, symplekin, and small rab family GTPases. The first identified component of the tight junction strands, which form the actual junction between cells, was the integral membrane protein occludin, a 65 kD protein with four transmembrane domains. ZO-1 binds to the carboxy-terminal region of occludin and may localize occludin to the tight junction. A recently identified family of proteins, the claudins, are also components of tight junction strands.

Claudins are 22-25 kD proteins which also contain four conserved transmembrane domains, but have no sequence homology to occludin (Furuse, M. et al. (1998) J. Cell Biol. 141:1539-1550). At least eight members of the claudin family have been cloned from mice. Claudin-2, in particular, is the most ancestral member of the claudin family. The eight claudins studied so far all have distinct tissue distributions and claudin-6, in particular, appears to be developmentally regulated (Morita, K. et al. (1999) Proc. Natl. Acad. Sci. USA 96:511-516). Claudin-2 expression is primarily restricted to the liver and kidney, with low levels of expression in the brain (Furuse et al. <u>supra</u>). Both claudin-1 and claudin-2 localize exclusively to the tight junction by immunofluorescence, and introduction of cDNA for claudin-1 and claudin-2 into mouse fibroblasts lacking tight junctions could induce tight junction formation (Furuse, M. et al. (1998) J. Cell Biol. 143:391-401). As occludin expression induces only a small number of short

strands, the claudins appear to be the major structural components of tight junction strands, with occludin being an accessory protein.

Other classes of transmembrane proteins involved in different types of cell junction formation and cell adhesion are the integrins, cadherins, and selectins. Integrins are transmembrane receptors at focal adhesions, actin-based cell junctions that occur between cells and the extracellular matrix. It now appears that a second class of cell surface molecules modify the type of adhesion mediated by primary integrin receptors. In particular, the syndecans, a family of heparan sulfate proteoglycans, act as co-receptors in adhesion and modify the downstream organization of the cytoskeleton. For example, although integrin is sufficient for attachment and spreading of primary fibroblasts, a secondary signal through interactions of matrix molecules with syndecans is needed for later stages of focal adhesion and stress fiber formation. The glycosaminoglycan chains of syndecans interact with the heparan binding domains of matrix proteins such as fibronectin, laminin, tenascin, and collagens, as well as with growth factors, proteases and protease inhibitors. Syndecans also interact with components of downstream signaling pathways, including protein kinase C and the src/cortactin pathway (Woods, A. et al. (1998) Matrix Biol. 17:477-483; Rapraeger, A.C. and V.L. Ott (1998) Curr. Opin. Cell Biol. 10:620-628).

A recently discovered protein, syntenin, interacts with the cytoplasmic domains of syndecans and may form the link between syndecans and the cytoskeleton (Grootjans, J.J. et al. (1997) Proc. Natl. Acad. Sci. USA 94:13683-13688). Syntenin contains two tandem PDZ domains (from K110 through P193 and from F194 through I274 in GI 3342560; SEQ ID NO:12), both of which are required to bind syndecan. Syntenin coclusters with syndecans at the plasma membrane, and its localization is affected by overexpression of syndecans 1, 2, and 4. Overexpression of syntenin results in cells that are larger, flatter, and have many cell surface projections, demonstrating the effect of syntenin on membrane dynamics and microfilament organization (Grootjans et al. supra).

The discovery of new human membrane-associated organizational proteins and the polynucleotides encoding them satisfies a need in the art by providing new compositions which are useful in the diagnosis, prevention, and treatment of cell proliferative disorders, including cancer, and autoimmune/inflammatory, neurological, developmental, vesicle trafficking, reproductive, gastrointestinal, and renal disorders.

SUMMARY OF THE INVENTION

The invention features substantially purified polypeptides, human membrane-associated organizational proteins, referred to collectively as "HJNCT" and individually as "HJNCT-1," "HJNCT-2," "HJNCT-3," and "HJNCT-4." In one aspect, the invention provides a substantially purified polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-4,

and fragments thereof.

The invention further provides a substantially purified variant having at least 95% amino acid identity to at least one of the amino acid sequences selected from the group consisting of SEQ ID NO:1-4, and fragments thereof. The invention also provides an isolated and purified polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-4, and fragments thereof. The invention also includes an isolated and purified polynucleotide variant having at least 90% polynucleotide sequence identity to the polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-4, and fragments thereof.

Additionally, the invention provides an isolated and purified polynucleotide which hybridizes under stringent conditions to the polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-4, and fragments thereof. The invention also provides an isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide encoding the polypeptide comprising the amino acid sequence selected from the group consisting of SEQ ID NO:1-4, and fragments thereof.

The invention also provides a method for detecting a polynucleotide in a sample containing nucleic acids, the method comprising the steps of (a) hybridizing the complement of the polynucleotide sequence to at least one of the polynucleotides of the sample, thereby forming a hybridization complex; and (b) detecting the hybridization complex, wherein the presence of the hybridization complex correlates with the presence of a polynucleotide in the sample. In one aspect, the method further comprises amplifying the polynucleotide prior to hybridization.

The invention also provides an isolated and purified polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:5-8, and fragments thereof. The invention further provides an isolated and purified polynucleotide variant having at least 90% polynucleotide sequence identity to the polynucleotide sequence selected from the group consisting of SEQ ID NO:5-8, and fragments thereof. The invention also provides an isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:5-8, and fragments thereof.

The invention further provides an expression vector containing at least a fragment of the polynucleotide encoding the polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-4, and fragments thereof. In another aspect, the expression vector is contained within a host cell.

The invention also provides a method for producing a polypeptide, the method comprising the steps of: (a) culturing the host cell containing an expression vector containing at least a fragment of a polynucleotide under conditions suitable for the expression of the polypeptide; and (b) recovering the

polypeptide from the host cell culture.

The invention also provides a pharmaceutical composition comprising a substantially purified polypeptide having the amino acid sequence selected from the group consisting of SEQ ID NO:1-4, and fragments thereof, in conjunction with a suitable pharmaceutical carrier.

The invention further includes a purified antibody which binds to a polypeptide selected from the group consisting of SEQ ID NO:1-4, and fragments thereof. The invention also provides a purified agonist and a purified antagonist to the polypeptide.

The invention also provides a method for treating or preventing a disorder associated with decreased expression or activity of HJNCT, the method comprising administering to a subject in need of such treatment an effective amount of a pharmaceutical composition comprising a substantially purified polypeptide having the amino acid sequence selected from the group consisting of SEQ ID NO:1-4, and fragments thereof, in conjunction with a suitable pharmaceutical carrier.

The invention also provides a method for treating or preventing a disorder associated with increased expression or activity of HJNCT, the method comprising administering to a subject in need of such treatment an effective amount of an antagonist of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-4, and fragments thereof.

BRIEF DESCRIPTION OF THE FIGURES AND TABLES

Figures 1A, 1B, and 1C show the amino acid sequence alignment between HJNCT-1 (Incyte Clone number 2687924; SEQ ID NO:1) and rat Homer (GI 1913909; SEQ ID NO:9), produced using the multisequence alignment program of LASERGENE software (DNASTAR Inc, Madison WI).

Figure 2 shows the amino acid sequence alignment between HJNCT-3 (Incyte Clone number 2594049; SEQ ID NO:3) and mouse claudin-2 (GI 3335184; SEQ ID NO:11), produced using the multisequence alignment program of LASERGENE software.

Figures 3A and 3B show the amino acid sequence alignment between HJNCT-4 (Incyte Clone number 5139028; SEQ ID NO:4) and mouse syntenin (GI 3342560; SEQ ID NO:12), produced using the multisequence alignment program of LASERGENE software.

Table 1 shows polypeptide and nucleotide sequence identification numbers (SEQ ID NOs), clone identification numbers (clone IDs), cDNA libraries, and cDNA fragments used to assemble full-length sequences encoding HJNCT.

Table 2 shows features of each polypeptide sequence, including potential motifs, homologous sequences, and methods and algorithms used for identification of HJNCT.

Table 3 shows useful fragments of each nucleic acid sequence; the tissue-specific expression patterns of each nucleic acid sequence as determined by northern analysis; diseases, disorders, or

conditions associated with these tissues; and the vector into which each cDNA was cloned.

Table 4 describes the tissues used to construct the cDNA libraries from which cDNA clones encoding HJNCT were isolated.

Table 5 shows the tools, programs, and algorithms used to analyze HJNCT, along with applicable descriptions, references, and threshold parameters.

DESCRIPTION OF THE INVENTION

Before the present proteins, nucleotide sequences, and methods are described, it is understood that this invention is not limited to the particular machines, materials and methods described, as these may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention which will be limited only by the appended claims.

It must be noted that as used herein and in the appended claims, the singular forms "a," "an," and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, a reference to "a host cell" includes a plurality of such host cells, and a reference to "an antibody" is a reference to one or more antibodies and equivalents thereof known to those skilled in the art, and so forth.

Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any machines, materials, and methods similar or equivalent to those described herein can be used to practice or test the present invention, the preferred machines, materials and methods are now described. All publications mentioned herein are cited for the purpose of describing and disclosing the cell lines, protocols, reagents and vectors which are reported in the publications and which might be used in connection with the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

DEFINITIONS

"HJNCT" refers to the amino acid sequences of substantially purified HJNCT obtained from any species, particularly a mammalian species, including bovine, ovine, porcine, murine, equine, and preferably the human species, from any source, whether natural, synthetic, semi-synthetic, or recombinant.

The term "agonist" refers to a molecule which, when bound to HJNCT, increases or prolongs the duration of the effect of HJNCT. Agonists may include proteins, nucleic acids, carbohydrates, or any other molecules which bind to and modulate the effect of HJNCT.

An "allelic variant" is an alternative form of the gene encoding HJNCT. Allelic variants may result from at least one mutation in the nucleic acid sequence and may result in altered mRNAs or in

polypeptides whose structure or function may or may not be altered. Any given natural or recombinant gene may have none, one, or many allelic forms. Common mutational changes which give rise to allelic variants are generally ascribed to natural deletions, additions, or substitutions of nucleotides. Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

"Altered" nucleic acid sequences encoding HJNCT include those sequences with deletions, insertions, or substitutions of different nucleotides, resulting in a polynucleotide the same as HJNCT or a polypeptide with at least one functional characteristic of HJNCT. Included within this definition are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding HJNCT, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide sequence encoding HJNCT. The encoded protein may also be "altered," and may contain deletions, insertions, or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent HJNCT. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues, as long as the biological or immunological activity of HJNCT is retained. For example, negatively charged amino acids may include aspartic acid and glutamic acid, positively charged amino acids may include lysine and arginine, and amino acids with uncharged polar head groups having similar hydrophilicity values may include leucine, isoleucine, and valine; glycine and alanine; asparagine and glutamine; serine and threonine; and phenylalanine and tyrosine.

The terms "amino acid" and "amino acid sequence" refer to an oligopeptide, peptide, polypeptide, or protein sequence, or a fragment of any of these, and to naturally occurring or synthetic molecules. In this context, "fragments," "immunogenic fragments," or "antigenic fragments" refer to fragments of HJNCT which are preferably at least 5 to about 15 amino acids in length, most preferably at least 14 amino acids, and which retain some biological activity or immunological activity of HJNCT. Where "amino acid sequence" is recited to refer to an amino acid sequence of a naturally occurring protein molecule, "amino acid sequence" and like terms are not meant to limit the amino acid sequence to the complete native amino acid sequence associated with the recited protein molecule.

"Amplification" relates to the production of additional copies of a nucleic acid sequence. Amplification is generally carried out using polymerase chain reaction (PCR) technologies well known in the art.

The term "antagonist" refers to a molecule which, when bound to HJNCT, decreases the amount or the duration of the effect of the biological or immunological activity of HJNCT. Antagonists may include proteins, nucleic acids, carbohydrates, antibodies, or any other molecules which decrease the

effect of HJNCT.

The term "antibody" refers to intact molecules as well as to fragments thereof, such as Fab, $F(ab')_2$, and Fv fragments, which are capable of binding the epitopic determinant. Antibodies that bind HJNCT polypeptides can be prepared using intact polypeptides or using fragments containing small peptides of interest as the immunizing antigen. The polypeptide or oligopeptide used to immunize an animal (e.g., a mouse, a rat, or a rabbit) can be derived from the translation of RNA, or synthesized chemically, and can be conjugated to a carrier protein if desired. Commonly used carriers that are chemically coupled to peptides include bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin (KLH). The coupled peptide is then used to immunize the animal.

The term "antigenic determinant" refers to that fragment of a molecule (i.e., an epitope) that makes contact with a particular antibody. When a protein or a fragment of a protein is used to immunize a host animal, numerous regions of the protein may induce the production of antibodies which bind specifically to antigenic determinants (given regions or three-dimensional structures on the protein). An antigenic determinant may compete with the intact antigen (i.e., the immunogen used to elicit the immune response) for binding to an antibody.

The term "antisense" refers to any composition containing a nucleic acid sequence which is complementary to the "sense" strand of a specific nucleic acid sequence. Antisense molecules may be produced by any method including synthesis or transcription. Once introduced into a cell, the complementary nucleotides combine with natural sequences produced by the cell to form duplexes and to block either transcription or translation. The designation "negative" can refer to the antisense strand, and the designation "positive" can refer to the sense strand.

The term "biologically active" refers to a protein having structural, regulatory, or biochemical functions of a naturally occurring molecule. Likewise, "immunologically active" refers to the capability of the natural, recombinant, or synthetic HJNCT, or of any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

The terms "complementary" and "complementarity" refer to the natural binding of polynucleotides by base pairing. For example, the sequence "5' A-G-T 3'" bonds to the complementary sequence "3' T-C-A 5'." Complementarity between two single-stranded molecules may be "partial," such that only some of the nucleic acids bind, or it may be "complete," such that total complementarity exists between the single stranded molecules. The degree of complementarity between nucleic acid strands has significant effects on the efficiency and strength of the hybridization between the nucleic acid strands. This is of particular importance in amplification reactions, which depend upon binding between nucleic acids strands, and in the design and use of peptide nucleic acid (PNA) molecules.

A "composition comprising a given polynucleotide sequence" and a "composition comprising a

given amino acid sequence" refer broadly to any composition containing the given polynucleotide or amino acid sequence. The composition may comprise a dry formulation or an aqueous solution. Compositions comprising polynucleotide sequences encoding HJNCT or fragments of HJNCT may be employed as hybridization probes. The probes may be stored in freeze-dried form and may be associated with a stabilizing agent such as a carbohydrate. In hybridizations, the probe may be deployed in an aqueous solution containing salts (e.g., NaCl), detergents (e.g., sodium dodecyl sulfate; SDS), and other components (e.g., Denhardt's solution, dry milk, salmon sperm DNA, etc.).

"Consensus sequence" refers to a nucleic acid sequence which has been resequenced to resolve uncalled bases, extended using the XL-PCR kit (Perkin-Elmer, Norwalk CT) in the 5' and/or the 3' direction, and resequenced, or which has been assembled from the overlapping sequences of more than one Incyte Clone using a computer program for fragment assembly, such as the GELVIEW fragment assembly system (GCG, Madison WI). Some sequences have been both extended and assembled to produce the consensus sequence.

The term "correlates with expression of a polynucleotide" indicates that the detection of the presence of nucleic acids, the same or related to a nucleic acid sequence encoding HJNCT, by northern analysis is indicative of the presence of nucleic acids encoding HJNCT in a sample, and thereby correlates with expression of the transcript from the polynucleotide encoding HJNCT.

A "deletion" refers to a change in the amino acid or nucleotide sequence that results in the absence of one or more amino acid residues or nucleotides.

The term "derivative" refers to the chemical modification of a polypeptide sequence, or a polynucleotide sequence. Chemical modifications of a polynucleotide sequence can include, for example, replacement of hydrogen by an alkyl, acyl, or amino group. A derivative polynucleotide encodes a polypeptide which retains at least one biological or immunological function of the natural molecule. A derivative polypeptide is one modified by glycosylation, pegylation, or any similar process that retains at least one biological or immunological function of the polypeptide from which it was derived.

The term "similarity" refers to a degree of complementarity. There may be partial similarity or complete similarity. The word "identity" may substitute for the word "similarity." A partially complementary sequence that at least partially inhibits an identical sequence from hybridizing to a target nucleic acid is referred to as "substantially similar." The inhibition of hybridization of the completely complementary sequence to the target sequence may be examined using a hybridization assay (Southern or northern blot, solution hybridization, and the like) under conditions of reduced stringency. A substantially similar sequence or hybridization probe will compete for and inhibit the binding of a completely similar (identical) sequence to the target sequence under conditions of reduced stringency. This is not to say that conditions of reduced stringency are such that non-specific binding is permitted, as

reduced stringency conditions require that the binding of two sequences to one another be a specific (i.e., a selective) interaction. The absence of non-specific binding may be tested by the use of a second target sequence which lacks even a partial degree of complementarity (e.g., less than about 30% similarity or identity). In the absence of non-specific binding, the substantially similar sequence or probe will not hybridize to the second non-complementary target sequence.

The phrases "percent identity" and "% identity" refer to the percentage of sequence similarity found in a comparison of two or more amino acid or nucleic acid sequences. Percent identity can be determined electronically, e.g., by using the MEGALIGN program (DNASTAR, Madison WI) which creates alignments between two or more sequences according to methods selected by the user, e.g., the clustal method. (See, e.g., Higgins, D.G. and P.M. Sharp (1988) Gene 73:237-244.) Parameters for each method may be the default parameters provided by MEGALIGN or may be specified by the user. The clustal algorithm groups sequences into clusters by examining the distances between all pairs. The clusters are aligned pairwise and then in groups. The percentage similarity between two amino acid sequences, e.g., sequence A and sequence B, is calculated by dividing the length of sequence A, minus the number of gap residues in sequence A, minus the number of gap residues in sequence B, into the sum of the residue matches between sequence A and sequence B, times one hundred. Gaps of low or of no similarity between the two amino acid sequences are not included in determining percentage similarity. Percent identity between nucleic acid sequences can also be counted or calculated by other methods known in the art, e.g., the Jotun Hein method. (See, e.g., Hein, J. (1990) Methods Enzymol. 183:626-645.) Identity between sequences can also be determined by other methods known in the art, e.g., by varying hybridization conditions.

"Human artificial chromosomes" (HACs) are linear microchromosomes which may contain DNA sequences of about 6 kb to 10 Mb in size, and which contain all of the elements required for stable mitotic chromosome segregation and maintenance.

The term "humanized antibody" refers to antibody molecules in which the amino acid sequence in the non-antigen binding regions has been altered so that the antibody more closely resembles a human antibody, and still retains its original binding ability.

"Hybridization" refers to any process by which a strand of nucleic acid binds with a complementary strand through base pairing.

The term "hybridization complex" refers to a complex formed between two nucleic acid sequences by virtue of the formation of hydrogen bonds between complementary bases. A hybridization complex may be formed in solution (e.g., C_0t or R_0t analysis) or formed between one nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate substrate to which cells or

their nucleic acids have been fixed).

The words "insertion" and "addition" refer to changes in an amino acid or nucleotide sequence resulting in the addition of one or more amino acid residues or nucleotides, respectively, to the sequence found in the naturally occurring molecule.

"Immune response" can refer to conditions associated with inflammation, trauma, immune disorders, or infectious or genetic disease, etc. These conditions can be characterized by expression of various factors, e.g., cytokines, chemokines, and other signaling molecules, which may affect cellular and systemic defense systems.

The term "microarray" refers to an arrangement of distinct polynucleotides on a substrate.

The terms "element" and "array element" in a microarray context, refer to hybridizable polynucleotides arranged on the surface of a substrate.

The term "modulate" refers to a change in the activity of HJNCT. For example, modulation may cause an increase or a decrease in protein activity, binding characteristics, or any other biological, functional, or immunological properties of HJNCT.

The phrases "nucleic acid" or "nucleic acid sequence," as used herein, refer to a nucleotide, oligonucleotide, polynucleotide, or any fragment thereof. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA), or to any DNA-like or RNA-like material. In this context, "fragments" refers to those nucleic acid sequences which comprise a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:5-8, for example, as distinct from any other sequence in the same genome. For example, a fragment of SEQ ID NO:5-8 is useful in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:5-8 from related polynucleotide sequences. A fragment of SEQ ID NO:5-8 is at least about 15-20 nucleotides in length. The precise length of the fragment of SEQ ID NO:5-8 and the region of SEQ ID NO:5-8 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment. In some cases, a fragment, when translated, would produce polypeptides retaining some functional characteristic, e.g., antigenicity, or structural domain characteristic, e.g., ATP-binding site, of the full-length polypeptide.

The terms "operably associated" and "operably linked" refer to functionally related nucleic acid sequences. A promoter is operably associated or operably linked with a coding sequence if the promoter controls the translation of the encoded polypeptide. While operably associated or operably linked nucleic acid sequences can be contiguous and in the same reading frame, certain genetic elements, e.g., repressor genes, are not contiguously linked to the sequence encoding the polypeptide but still bind to operator sequences that control expression of the polypeptide.

The term "oligonucleotide" refers to a nucleic acid sequence of at least about 6 nucleotides to 60 nucleotides, preferably about 15 to 30 nucleotides, and most preferably about 20 to 25 nucleotides, which can be used in PCR amplification or in a hybridization assay or microarray. "Oligonucleotide" is substantially equivalent to the terms "amplimer," "primer," "oligomer," and "probe," as these terms are commonly defined in the art.

"Peptide nucleic acid" (PNA) refers to an antisense molecule or anti-gene agent which comprises an oligonucleotide of at least about 5 nucleotides in length linked to a peptide backbone of amino acid residues ending in lysine. The terminal lysine confers solubility to the composition. PNAs preferentially bind complementary single stranded DNA or RNA and stop transcript elongation, and may be pegylated to extend their lifespan in the cell.

The term "sample" is used in its broadest sense. A sample suspected of containing nucleic acids encoding HJNCT, or fragments thereof, or HJNCT itself, may comprise a bodily fluid; an extract from a cell, chromosome, organelle, or membrane isolated from a cell; a cell; genomic DNA, RNA, or cDNA, in solution or bound to a substrate; a tissue; a tissue print; etc.

The terms "specific binding" and "specifically binding" refer to that interaction between a protein or peptide and an agonist, an antibody, or an antagonist. The interaction is dependent upon the presence of a particular structure of the protein, e.g., the antigenic determinant or epitope, recognized by the binding molecule. For example, if an antibody is specific for epitope "A," the presence of a polypeptide containing the epitope A, or the presence of free unlabeled A, in a reaction containing free labeled A and the antibody will reduce the amount of labeled A that binds to the antibody.

The term "stringent conditions" refers to conditions which permit hybridization between polynucleotides and the claimed polynucleotides. Stringent conditions can be defined by salt concentration, the concentration of organic solvent, e.g., formamide, temperature, and other conditions well known in the art. In particular, stringency can be increased by reducing the concentration of salt, increasing the concentration of formamide, or raising the hybridization temperature.

The term "substantially purified" refers to nucleic acid or amino acid sequences that are removed from their natural environment and are isolated or separated, and are at least about 60% free, preferably about 75% free, and most preferably about 90% free from other components with which they are naturally associated.

A "substitution" refers to the replacement of one or more amino acids or nucleotides by different amino acids or nucleotides, respectively.

"Substrate" refers to any suitable rigid or semi-rigid support including membranes, filters, chips, slides, wafers, fibers, magnetic or nonmagnetic beads, gels, tubing, plates, polymers, microparticles and capillaries. The substrate can have a variety of surface forms, such as wells, trenches, pins, channels and

pores, to which polynucleotides or polypeptides are bound.

"Transformation" describes a process by which exogenous DNA enters and changes a recipient cell. Transformation may occur under natural or artificial conditions according to various methods well known in the art, and may rely on any known method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. The method for transformation is selected based on the type of host cell being transformed and may include, but is not limited to, viral infection, electroporation, heat shock, lipofection, and particle bombardment. The term "transformed" cells includes stably transformed cells in which the inserted DNA is capable of replication either as an autonomously replicating plasmid or as part of the host chromosome, as well as transiently transformed cells which express the inserted DNA or RNA for limited periods of time.

A "variant" of HJNCT polypeptides refers to an amino acid sequence that is altered by one or more amino acid residues. The variant may have "conservative" changes, wherein a substituted amino acid has similar structural or chemical properties (e.g., replacement of leucine with isoleucine). More rarely, a variant may have "nonconservative" changes (e.g., replacement of glycine with tryptophan). Analogous minor variations may also include amino acid deletions or insertions, or both. Guidance in determining which amino acid residues may be substituted, inserted, or deleted without abolishing biological or immunological activity may be found using computer programs well known in the art, for example, LASERGENE software (DNASTAR).

The term "variant," when used in the context of a polynucleotide sequence, may encompass a polynucleotide sequence related to HJNCT. This definition may also include, for example, "allelic" (as defined above), "splice," "species," or "polymorphic" variants. A splice variant may have significant identity to a reference molecule, but will generally have a greater or lesser number of polynucleotides due to alternate splicing of exons during mRNA processing. The corresponding polypeptide may possess additional functional domains or an absence of domains. Species variants are polynucleotide sequences that vary from one species to another. The resulting polypeptides generally will have significant amino acid identity relative to each other. A polymorphic variant is a variation in the polynucleotide sequence of a particular gene between individuals of a given species. Polymorphic variants also may encompass "single nucleotide polymorphisms" (SNPs) in which the polynucleotide sequence varies by one base. The presence of SNPs may be indicative of, for example, a certain population, a disease state, or a propensity for a disease state.

THE INVENTION

The invention is based on the discovery of new human membrane-associated organizational proteins (HJNCT), the polynucleotides encoding HJNCT, and the use of these compositions for the diagnosis, treatment, or prevention of cell proliferative disorders, including cancer, and

autoimmune/inflammatory, neurological, developmental, vesicle trafficking, reproductive, gastrointestinal, and renal disorders.

Table 1 lists the Incyte clones used to assemble full length nucleotide sequences encoding HJNCT. Columns 1 and 2 show the sequence identification numbers (SEQ ID NOs) of the polypeptide and nucleotide sequences, respectively. Column 3 shows the clone IDs of the Incyte clones in which nucleic acids encoding each HJNCT were identified, and column 4 shows the cDNA libraries from which these clones were isolated. Column 5 shows Incyte clones and their corresponding cDNA libraries. Clones for which cDNA libraries are not indicated were derived from pooled cDNA libraries. The clones in column 5 were used to assemble the consensus nucleotide sequence of each HJNCT and are useful as fragments in hybridization technologies.

The columns of Table 2 show various properties of each of the polypeptides of the invention: column 1 references the SEQ ID NO; column 2 shows the number of amino acid residues in each polypeptide; column 3 shows potential phosphorylation sites; column 4 shows potential glycosylation sites; column 5 shows the amino acid residues comprising signature sequences and motifs; column 6 shows homologous sequences as identified by BLAST analysis; and column 7 shows analytical methods used to characterize each polypeptide through sequence homology and protein motifs.

As shown in Figures 1A, 1B, and 1C, HJNCT-1 has chemical and structural similarity with rat Homer (GI 1913909; SEQ ID NO:9). In particular, HJNCT-1 and rat Homer share 72% identity. Furthermore, HJNCT-1 and rat Homer share 100% identity within the region from residue T65 to F93, as shown in Figure 1A. Within this region, HJNCT-1 contains a putative GLGF segment from G90 to F93 and an arginine residue at R84. HJNCT-2 has chemical and structural similarity with rat GRASP65 (GI 4432587; SEQ ID NO:10). In particular, HJNCT-2 and rat GRASP65 share 71% identity. HJNCT-2 and rat GRASP65 are 100% identical from residue L50 to residue I62 of HJNCT-2, the region shown to be important for binding GM130 in GRASP65. Like rat GRASP65, HJNCT-2 contains many possible phosphorylation sites and has possible N-myristoylation sites. As shown in Figure 2, HJNCT-3 has chemical and structural homology with mouse claudin-2 (GI 3335184; SEQ ID NO:11). In particular, HJNCT-3 and mouse claudin-2 share 91% identity. As shown in Figures 3A and 3B, HJNCT-4 has chemical and structural homology with mouse syntenin (GI 3342560; SEQ ID NO:12). In particular, HJNCT-4 and mouse syntenin share 61% identity.

The columns of Table 3 show the tissue-specificity and diseases, disorders, or conditions associated with nucleotide sequences encoding HJNCT. The first column of Table 3 lists the nucleotide SEQ ID NOs. Column 2 lists fragments of the nucleotide sequences of column 1. These fragments are useful, for example, in hybridization or amplification technologies to identify SEQ ID NO:5-8 and to distinguish between SEQ ID NO:5-8 and related polynucleotide sequences. The polypeptides encoded by

these fragments are useful, for example, as immunogenic peptides. Column 3 lists tissue categories which express HJNCT as a fraction of total tissues expressing HJNCT. Column 4 lists diseases, disorders, or conditions associated with those tissues expressing HJNCT as a fraction of total tissues expressing HJNCT. Of particular note is the expression of HJNCT-1 in reproductive tissue, nervous tissue, and fetal cell lines; the expression of HJNCT-2 in tissues associated with cancer and cell proliferation disorders; the expression of HJNCT-3 in cancerous, fetal, proliferating, gastrointestinal, and urologic tissues; and the expression of HJNCT-4 in cancerous, fetal, proliferating, and gastrointestinal tissues. Column 5 lists the vectors used to subclone each cDNA library.

The columns of Table 4 show descriptions of the tissues used to construct the cDNA libraries from which cDNA clones encoding HJNCT were isolated. Column 1 references the nucleotide SEQ ID NOs, column 2 shows the cDNA libraries from which these clones were isolated, and column 3 shows the tissue origins and other descriptive information relevant to the cDNA libraries in column 2.

The invention also encompasses HJNCT variants. A preferred HJNCT variant is one which has at least about 85%, more preferably at least about 90%, and most preferably at least about 95% amino acid sequence identity to the HJNCT amino acid sequence, and which contains at least one functional or structural characteristic of HJNCT.

The invention also encompasses polynucleotides which encode HJNCT. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:5-8, which encodes HJNCT.

The invention also encompasses a variant of a polynucleotide sequence encoding HJNCT. In particular, such a variant polynucleotide sequence will have at least about 85%, more preferably at least about 90%, and most preferably at least about 95% polynucleotide sequence identity to the polynucleotide sequence encoding HJNCT. A particular aspect of the invention encompasses a variant of a sequence selected from the group consisting of SEQ ID NO:5-8 which has at least about 85%, more preferably at least about 90%, and most preferably at least about 95% polynucleotide sequence identity to a sequence selected from the group consisting of SEQ ID NO:5-8. Any one of the polynucleotide variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of HJCNT.

It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of polynucleotide sequences encoding HJNCT, some bearing minimal similarity to the polynucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of polynucleotide sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the polynucleotide sequence of naturally occurring

HJNCT, and all such variations are to be considered as being specifically disclosed.

Although nucleotide sequences which encode HJNCT and its variants are preferably capable of hybridizing to the nucleotide sequence of the naturally occurring HJNCT under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding HJNCT or its derivatives possessing a substantially different codon usage, e.g., inclusion of non-naturally occurring codons. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding HJNCT and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

The invention also encompasses production of DNA sequences which encode HJNCT and HJNCT derivatives, or fragments thereof, entirely by synthetic chemistry. After production, the synthetic sequence may be inserted into any of the many available expression vectors and cell systems using reagents well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into a sequence encoding HJNCT or any fragment thereof.

Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed polynucleotide sequences, and, in particular, to those shown in SEQ ID NO:5-8 and fragments thereof under various conditions of stringency. (See, e.g., Wahl, G.M. and S.L. Berger (1987) Methods Enzymol. 152:399-407; Kimmel, A.R. (1987) Methods Enzymol. 152:507-511.) For example, stringent salt concentration will ordinarily be less than about 750 mM NaCl and 75 mM trisodium citrate, preferably less than about 500 mM NaCl and 50 mM trisodium citrate, and most preferably less than about 250 mM NaCl and 25 mM trisodium citrate. Low stringency hybridization can be obtained in the absence of organic solvent, e.g., formamide, while high stringency hybridization can be obtained in the presence of at least about 35% formamide, and most preferably at least about 50% formamide. Stringent temperature conditions will ordinarily include temperatures of at least about 30°C, more preferably of at least about 37°C, and most preferably of at least about 42°C. Varying additional parameters, such as hybridization time, the concentration of detergent, e.g., sodium dodecyl sulfate (SDS), and the inclusion or exclusion of carrier DNA, are well known to those skilled in the art. Various levels of stringency are accomplished by combining these various conditions as needed. In a preferred embodiment, hybridization will occur at 30°C in 750 mM NaCl, 75 mM trisodium citrate, and 1% SDS. In a more preferred embodiment, hybridization will occur at 37°C in 500 mM NaCl, 50 mM trisodium citrate, 1% SDS, 35% formamide, and 100 μg/ml denatured salmon sperm DNA (ssDNA). In a most preferred embodiment, hybridization will occur at 42°C in 250 mM NaCl, 25 mM trisodium citrate, 1% SDS, 50 % formamide,

and 200 μg/ml ssDNA. Useful variations on these conditions will be readily apparent to those skilled in the art.

The washing steps which follow hybridization can also vary in stringency. Wash stringency conditions can be defined by salt concentration and by temperature. As above, wash stringency can be increased by decreasing salt concentration or by increasing temperature. For example, stringent salt concentration for the wash steps will preferably be less than about 30 mM NaCl and 3 mM trisodium citrate, and most preferably less than about 15 mM NaCl and 1.5 mM trisodium citrate. Stringent temperature conditions for the wash steps will ordinarily include temperature of at least about 25°C, more preferably of at least about 42°C, and most preferably of at least about 68°C. In a preferred embodiment, wash steps will occur at 25°C in 30 mM NaCl, 3 mM trisodium citrate, and 0.1% SDS. In a more preferred embodiment, wash steps will occur at 42°C in 15 mM NaCl, 1.5 mM trisodium citrate, and 0.1% SDS. In a most preferred embodiment, wash steps will occur at 68°C in 15 mM NaCl, 1.5 mM trisodium citrate, and 0.1% SDS. Additional variations on these conditions will be readily apparent to those skilled in the art.

Methods for DNA sequencing are well known in the art and may be used to practice any of the embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, SEQUENASE (US Biochemical, Cleveland OH), Taq polymerase (Perkin-Elmer), thermostable T7 polymerase (Amersham Pharmacia Biotech, Piscataway NJ), or combinations of polymerases and proofreading exonucleases such as those found in the ELONGASE amplification system (Life Technologies, Gaithersburg MD). Preferably, sequence preparation is automated with machines such as the MICROLAB 2200 liquid transfer system (Hamilton, Reno NV), PTC200 thermal cycler (MJ Research, Watertown MA) and ABI CATALYST 800 thermal cycler (Perkin-Elmer). Sequencing is then carried out using either the ABI 373 or 377 DNA sequencing system (Perkin-Elmer), the MEGABACE 1000 DNA sequencing system (Molecular Dynamics, Sunnyvale CA), or other systems known in the art. The resulting sequences are analyzed using a variety of algorithms which are well known in the art. (See, e.g., Ausubel, F.M. (1997) Short Protocols in Molecular Biology, John Wiley & Sons, New York NY, unit 7.7; Meyers, R.A. (1995) Molecular Biology and Biotechnology, Wiley VCH, New York NY, pp. 856-853.)

The nucleic acid sequences encoding HJNCT may be extended utilizing a partial nucleotide sequence and employing various PCR-based methods known in the art to detect upstream sequences, such as promoters and regulatory elements. For example, one method which may be employed, restriction-site PCR, uses universal and nested primers to amplify unknown sequence from genomic DNA within a cloning vector. (See, e.g., Sarkar, G. (1993) PCR Methods Applic. 2:318-322.) Another method, inverse PCR, uses primers that extend in divergent directions to amplify unknown sequence from a circularized

template. The template is derived from restriction fragments comprising a known genomic locus and surrounding sequences. (See, e.g., Triglia, T. et al. (1988) Nucleic Acids Res. 16:8186.) A third method, capture PCR, involves PCR amplification of DNA fragments adjacent to known sequences in human and yeast artificial chromosome DNA. (See, e.g., Lagerstrom, M. et al. (1991) PCR Methods Applic. 1:111-119.) In this method, multiple restriction enzyme digestions and ligations may be used to insert an engineered double-stranded sequence into a region of unknown sequence before performing PCR. Other methods which may be used to retrieve unknown sequences are known in the art. (See, e.g., Parker, J.D. et al. (1991) Nucleic Acids Res. 19:3055-3060). Additionally, one may use PCR, nested primers, and PROMOTERFINDER libraries (Clontech, Palo Alto CA) to walk genomic DNA. This procedure avoids the need to screen libraries and is useful in finding intron/exon junctions. For all PCR-based methods, primers may be designed using commercially available software, such as OLIGO 4.06 primer analysis software (National Biosciences, Plymouth MN) or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the template at temperatures of about 68°C to 72°C.

When screening for full-length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. In addition, random-primed libraries, which often include sequences containing the 5' regions of genes, are preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries may be useful for extension of sequence into 5' non-transcribed regulatory regions.

Capillary electrophoresis systems which are commercially available may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products. In particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different nucleotide-specific, laser-stimulated fluorescent dyes, and a charge coupled device camera for detection of the emitted wavelengths. Output/light intensity may be converted to electrical signal using appropriate software (e.g., GENOTYPER and SEQUENCE NAVIGATOR, Perkin-Elmer), and the entire process from loading of samples to computer analysis and electronic data display may be computer controlled. Capillary electrophoresis is especially preferable for sequencing small DNA fragments which may be present in limited amounts in a particular sample.

In another embodiment of the invention, polynucleotide sequences or fragments thereof which encode HJNCT may be cloned in recombinant DNA molecules that direct expression of HJNCT, or fragments or functional equivalents thereof, in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be produced and used to express HJNCT.

The nucleotide sequences of the present invention can be engineered using methods generally

known in the art in order to alter HJNCT-encoding sequences for a variety of purposes including, but not limited to, modification of the cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. For example, oligonucleotide-mediated site-directed mutagenesis may be used to introduce mutations that create new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, and so forth.

In another embodiment, sequences encoding HJNCT may be synthesized, in whole or in part, using chemical methods well known in the art. (See, e.g., Caruthers, M.H. et al. (1980) Nucleic Acids Symp. Ser. 7:215-223, and Horn, T. et al. (1980) Nucleic Acids Symp. Ser. 7:225-232.) Alternatively, HJNCT itself or a fragment thereof may be synthesized using chemical methods. For example, peptide synthesis can be performed using various solid-phase techniques. (See, e.g., Roberge, J.Y. et al. (1995) Science 269:202-204.) Automated synthesis may be achieved using the ABI 431A peptide synthesizer (Perkin-Elmer). Additionally, the amino acid sequence of HJNCT, or any part thereof, may be altered during direct synthesis and/or combined with sequences from other proteins, or any part thereof, to produce a variant polypeptide.

The peptide may be substantially purified by preparative high performance liquid chromatography. (See, e.g, Chiez, R.M. and F.Z. Regnier (1990) Methods Enzymol. 182:392-421.) The composition of the synthetic peptides may be confirmed by amino acid analysis or by sequencing. (See, e.g., Creighton, T. (1984) Proteins, Structures and Molecular Properties, WH Freeman, New York NY.)

In order to express a biologically active HJNCT, the nucleotide sequences encoding HJNCT or derivatives thereof may be inserted into an appropriate expression vector, i.e., a vector which contains the necessary elements for transcriptional and translational control of the inserted coding sequence in a suitable host. These elements include regulatory sequences, such as enhancers, constitutive and inducible promoters, and 5' and 3' untranslated regions in the vector and in polynucleotide sequences encoding HJNCT. Such elements may vary in their strength and specificity. Specific initiation signals may also be used to achieve more efficient translation of sequences encoding HJNCT. Such signals include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where sequences encoding HJNCT and its initiation codon and upstream regulatory sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals including an in-frame ATG initiation codon should be provided by the vector. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers appropriate for the particular host cell system used. (See, e.g., Scharf, D. et al. (1994) Results Probl. Cell Differ. 20:125-162.)

Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding HJNCT and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. (See, e.g., Sambrook, J. et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview NY, ch. 4, 8, and 16-17; Ausubel, F.M. et al. (1995) Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, ch. 9, 13, and 16.)

A variety of expression vector/host systems may be utilized to contain and express sequences encoding HJNCT. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with viral expression vectors (e.g., baculovirus); plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV, or tobacco mosaic virus,TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems. The invention is not limited by the host cell employed.

In bacterial systems, a number of cloning and expression vectors may be selected depending upon the use intended for polynucleotide sequences encoding HJNCT. For example, routine cloning, subcloning, and propagation of polynucleotide sequences encoding HJNCT can be achieved using a multifunctional E. coli vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or PSPORT1 plasmid (Life Technologies). Ligation of sequences encoding HJNCT into the vector's multiple cloning site disrupts the *lacZ* gene, allowing a colorimetric screening procedure for identification of transformed bacteria containing recombinant molecules. In addition, these vectors may be useful for in vitro transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of nested deletions in the cloned sequence. (See, e.g., Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509.) When large quantities of HJNCT are needed, e.g. for the production of antibodies, vectors which direct high level expression of HJNCT may be used. For example, vectors containing the strong, inducible T5 or T7 bacteriophage promoter may be used.

Yeast expression systems may be used for production of HJNCT. A number of vectors containing constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH promoters, may be used in the yeast Saccharomyces cerevisiae or Pichia pastoris. In addition, such vectors direct either the secretion or intracellular retention of expressed proteins and enable integration of foreign sequences into the host genome for stable propagation. (See, e.g., Ausubel, 1995, supra; Bitter, G.A. et al. (1987) Methods Enzymol. 153:516-544; and Scorer, C.A. et al. (1994) Bio/Technology 12:181-184.)

Plant systems may also be used for expression of HJNCT. Transcription of sequences encoding HJNCT may be driven viral promoters, e.g., the 35S and 19S promoters of CaMV used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) EMBO J. 6:307-311).

Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used. (See, e.g., Coruzzi, G. et al. (1984) EMBO J. 3:1671-1680; Broglie, R. et al. (1984) Science 224:838-843; and Winter, J. et al. (1991) Results Probl. Cell Differ. 17:85-105.) These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. (See, e.g., The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196.)

In mammalian cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, sequences encoding HJNCT may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain infective virus which expresses HJNCT in host cells. (See, e.g., Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659.) In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells. SV40 or EBV-based vectors may also be used for high-level protein expression.

Human artificial chromosomes (HACs) may also be employed to deliver larger fragments of DNA than can be contained in and expressed from a plasmid. HACs of about 6 kb to 10 Mb are constructed and delivered via conventional delivery methods (liposomes, polycationic amino polymers, or vesicles) for therapeutic purposes. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355.)

For long term production of recombinant proteins in mammalian systems, stable expression of HJNCT in cell lines is preferred. For example, sequences encoding HJNCT can be transformed into cell lines using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for about 1 to 2 days in enriched media before being switched to selective media. The purpose of the selectable marker is to confer resistance to a selective agent, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be propagated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk*$^-$ or *apr*$^-$ cells, respectively. (See, e.g., Wigler, M. et al. (1977) Cell 11:223-232; Lowy, I. et al. (1980) Cell 22:817-823.) Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate; *neo* confers resistance to the aminoglycosides neomycin and G-418; and *als* and *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively. (See, e.g., Wigler, M. et al. (1980) Proc. Natl. Acad. Sci.

USA 77:3567-3570; Colbere-Garapin, F. et al. (1981) J. Mol. Biol. 150:1-14.) Additional selectable genes have been described, e.g., *trpB* and *hisD*, which alter cellular requirements for metabolites. (See, e.g., Hartman, S.C. and R.C. Mulligan (1988) Proc. Natl. Acad. Sci. USA 85:8047-8051.) Visible markers, e.g., anthocyanins, green fluorescent proteins (GFP; Clontech), ß glucuronidase and its substrate ß-glucuronide, or luciferase and its substrate luciferin may be used. These markers can be used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system. (See, e.g., Rhodes, C.A. (1995) Methods Mol. Biol. 55:121-131.)

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, the presence and expression of the gene may need to be confirmed. For example, if the sequence encoding HJNCT is inserted within a marker gene sequence, transformed cells containing sequences encoding HJNCT can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding HJNCT under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

In general, host cells that contain the nucleic acid sequence encoding HJNCT and that express HJNCT may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations, PCR amplification, and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein sequences.

Immunological methods for detecting and measuring the expression of HJNCT using either specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on HJNCT is preferred, but a competitive binding assay may be employed. These and other assays are well known in the art. (See, e.g., Hampton, R. et al. (1990) Serological Methods, a Laboratory Manual, APS Press, St. Paul MN, Sect. IV; Coligan, J.E. et al. (1997) Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York NY; and Pound, J.D. (1998) Immunochemical Protocols, Humana Press, Totowa NJ).

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding HJNCT include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, the sequences encoding HJNCT, or any fragments thereof, may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to

synthesize RNA probes in vitro by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits, such as those provided by Amersham Pharmacia Biotech, Promega (Madison WI), and US Biochemical. Suitable reporter molecules or labels which may be used for ease of detection include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with nucleotide sequences encoding HJNCT may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a transformed cell may be secreted or retained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode HJNCT may be designed to contain signal sequences which direct secretion of HJNCT through a prokaryotic or eukaryotic cell membrane.

In addition, a host cell strain may be chosen for its ability to modulate expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" form of the protein may also be used to specify protein targeting, folding, and/or activity. Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38) are available from the American Type Culture Collection (ATCC, Manassas VA) and may be chosen to ensure the correct modification and processing of the foreign protein.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences encoding HJNCT may be ligated to a heterologous sequence resulting in translation of a fusion protein in any of the aforementioned host systems. For example, a chimeric HJNCT protein containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of HJNCT activity. Heterologous protein and peptide moieties may also facilitate purification of fusion proteins using commercially available affinity matrices. Such moieties include, but are not limited to, glutathione S-transferase (GST), maltose binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG, *c-myc*, and hemagglutinin (HA). GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable immunoaffinity purification of fusion proteins using commercially available monoclonal and polyclonal antibodies that specifically recognize these epitope tags. A fusion protein may also be engineered to contain a proteolytic cleavage site located between the HJNCT

encoding sequence and the heterologous protein sequence, so that HJNCT may be cleaved away from the heterologous moiety following purification. Methods for fusion protein expression and purification are discussed in Ausubel (1995, supra, ch 10). A variety of commercially available kits may also be used to facilitate expression and purification of fusion proteins.

In a further embodiment of the invention, synthesis of radiolabeled HJNCT may be achieved in vitro using the TNT rabbit reticulocyte lysate or wheat germ extract systems (Promega). These systems couple transcription and translation of protein-coding sequences operably associated with the T7, T3, or SP6 promoters. Translation takes place in the presence of a radiolabeled amino acid precursor, preferably ^{35}S-methionine.

Fragments of HJNCT may be produced not only by recombinant production, but also by direct peptide synthesis using solid-phase techniques. (See, e.g., Creighton, supra, pp. 55-60.) Protein synthesis may be performed by manual techniques or by automation. Automated synthesis may be achieved, for example, using the ABI 431A peptide synthesizer (Perkin-Elmer). Various fragments of HJNCT may be synthesized separately and then combined to produce the full length molecule.

THERAPEUTICS

Chemical and structural similarity, e.g., in the context of sequences and motifs, exists between regions of HJNCT and human membrane-associated organizational proteins. In addition, the expression of HJNCT is closely associated with reproductive, neurological, developmental, cancerous, fetal or proliferating, gastrointestinal, and urologic tissues, and with inflammation and the immune response. Therefore, HJNCT appears to play a role in cell proliferative disorders, including cancer, and autoimmune/inflammatory, neurological, developmental, vesicle trafficking, reproductive, gastrointestinal, and renal disorders. In the treatment of disorders associated with increased HJNCT expression or activity, it is desirable to decrease the expression or activity of HJNCT. In the treatment of disorders associated with decreased HJNCT expression or activity, it is desirable to increase the expression or activity of HJNCT.

Therefore, in one embodiment, HJNCT or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of HJNCT. Examples of such disorders include, but are not limited to, a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancer, including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; an

autoimmune/inflammatory disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyenodocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis, inherited, metabolic, endocrine, and toxic myopathies, myasthenia gravis, periodic paralysis, mental disorders including mood, anxiety, and schizophrenic disorders, akathesia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, and Tourette's disorder; a developmental disorder such as renal tubular acidosis, anemia, Cushing's syndrome, achondroplastic dwarfism, Duchenne and Becker muscular dystrophy, epilepsy, gonadal dysgenesis, WAGR syndrome (Wilms' tumor, aniridia, genitourinary abnormalities, and mental retardation), Smith-Magenis syndrome, myelodysplastic syndrome, hereditary mucoepithelial dysplasia, hereditary keratodermas, hereditary neuropathies such as Charcot-Marie-Tooth disease and neurofibromatosis, hypothyroidism, hydrocephalus, seizure disorders such as Syndenham's chorea and cerebral palsy, spina

bifida, Williams syndrome, anencephaly, craniorachischisis, congenital glaucoma, cataract, sensorineural hearing loss, and any disorder associated with cell growth and differentiation, embryogenesis, and morphogenesis involving any tissue, organ, or system of a subject, e.g., the brain, adrenal gland, kidney, skeletal or reproductive system; a vesicle trafficking disorder such as cystic fibrosis, glucose-galactose malabsorption syndrome, hypercholesterolemia, diabetes mellitus, diabetes insipidus, hyper- and hypoglycemia, Grave's disease, goiter, Cushing's disease, and Addison's disease, other conditions associated with abnormal vesicle trafficking, including acquired immunodeficiency syndrome (AIDS), allergies including hay fever, asthma, and urticaria (hives); autoimmune hemolytic anemia, proliferative glomerulonephritis, inflammatory bowel disease, multiple sclerosis, myasthenia gravis, rheumatoid and osteoarthritis; scleroderma, Chediak-Higashi and Sjogren's syndromes; systemic lupus erythematosus, toxic shock syndrome, traumatic tissue damage, and viral, bacterial, fungal, helminthic, and protozoal infections; a reproductive disorder such as a disorder of prolactin production; infertility, including tubal disease, ovulatory defects, and endometriosis; disruptions of the estrous cycle, disruptions of the menstrual cycle, polycystic ovary syndrome, ovarian hyperstimulation syndrome, endometrial and ovarian tumors, uterine fibroids, autoimmune disorders, ectopic pregnancies, and teratogenesis; cancer of the breast, fibrocystic breast disease, and galactorrhea; disruptions of spermatogenesis, abnormal sperm physiology, cancer of the testis, cancer of the prostate, benign prostatic hyperplasia, prostatitis, Peyronie's disease, impotence, carcinoma of the male breast, and gynecomastia; a gastrointestinal disorder such as dysphagia, peptic esophagitis, esophageal spasm, esophageal stricture, esophageal carcinoma, dyspepsia, indigestion, gastritis, gastric carcinoma, anorexia, nausea, emesis, gastroparesis, antral or pyloric edema, abdominal angina, pyrosis, gastroenteritis, intestinal obstruction, infections of the intestinal tract, peptic ulcer, gastric ulcer, duodenal ulcer, cholelithiasis, cholecystitis, cholestasis, pancreatitis, pancreatic carcinoma, biliary tract disease, hepatitis, hyperbilirubinemia, cirrhosis, passive congestion of the liver, hepatoma, infectious colitis, ulcerative colitis, ulcerative proctitis, Crohn's disease, Whipple's disease, Mallory-Weiss syndrome, colonic carcinoma, colonic obstruction, irritable bowel syndrome, short bowel syndrome, diarrhea, constipation, gastrointestinal hemorrhage, acquired immunodeficiency syndrome (AIDS) enteropathy, jaundice, hepatic encephalopathy, hepatorenal syndrome, hepatic steatosis, hemochromatosis, Wilson's disease, alpha$_1$-antitrypsin deficiency, Reye's syndrome, primary sclerosing cholangitis, liver infarction, portal vein obstruction and thrombosis, centrilobular necrosis, peliosis hepatis, hepatic vein thrombosis, veno-occlusive disease, preeclampsia, eclampsia, acute fatty liver of pregnancy, intrahepatic cholestasis of pregnancy, and hepatic tumors including nodular hyperplasias, adenomas, and carcinomas; and a renal disorder such as renal amyloidosis, hypertension, primary aldosteronism, Addison's disease, renal failure, glomerulonephritis, chronic glomerulonephritis, tubulointerstitial nephritis, cystic disorders of the kidney and dysplastic malformations such as polycystic

disease, renal dysplasias, and cortical or medullary cysts, inherited polycystic renal diseases (PRD) such as recessive and autosomal dominant PRD, medullary cystic disease, medullary sponge kidney and tubular dysplasia, Alport's syndrome, non-renal cancers which affect renal physiology such as bronchogenic tumors of the lungs or tumors of the basal region of the brain, multiple myeloma, adenocarcinomas of the kidney, metastatic renal carcinoma, and nephrotoxic disorders including any functional or morphologic change in the kidney produced by any pharmaceutical, chemical, or biological agent that is ingested, injected, inhaled, or absorbed. Some broad categories of common nephrotoxic agents are heavy metals, all classes of antibiotics, analgesics, solvents, oxalosis-inducing agents, anticancer drugs, herbicides and pesticides, botanicals and biologicals, and antiepileptics.

In another embodiment, a vector capable of expressing HJNCT or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of HJNCT including, but not limited to, those described above.

In a further embodiment, a pharmaceutical composition comprising a substantially purified HJNCT in conjunction with a suitable pharmaceutical carrier may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of HJNCT including, but not limited to, those provided above.

In still another embodiment, an agonist which modulates the activity of HJNCT may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of HJNCT including, but not limited to, those listed above.

In a further embodiment, an antagonist of HJNCT may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of HJNCT. Examples of such disorders include, but are not limited to, those cell proliferative disorders, including cancer, and autoimmune/inflammatory, neurological, developmental, vesicle trafficking, reproductive, gastrointestinal, and renal disorders described above. In one aspect, an antibody which specifically binds HJNCT may be used directly as an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissue which express HJNCT.

In an additional embodiment, a vector expressing the complement of the polynucleotide encoding HJNCT may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of HJNCT including, but not limited to, those described above.

In other embodiments, any of the proteins, antagonists, antibodies, agonists, complementary sequences, or vectors of the invention may be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy may be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents may act synergistically to effect the treatment or prevention of the various disorders

described above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

An antagonist of HJNCT may be produced using methods which are generally known in the art. In particular, purified HJNCT may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind HJNCT. Antibodies to HJNCT may also be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, and single chain antibodies, Fab fragments, and fragments produced by a Fab expression library. Neutralizing antibodies (i.e., those which inhibit dimer formation) are especially preferred for therapeutic use.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others may be immunized by injection with HJNCT or with any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and Corynebacterium parvum are especially preferable.

It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to HJNCT have an amino acid sequence consisting of at least about 5 amino acids, and, more preferably, of at least about 10 amino acids. It is also preferable that these oligopeptides, peptides, or fragments are identical to a portion of the amino acid sequence of the natural protein and contain the entire amino acid sequence of a small, naturally occurring molecule. Short stretches of HJNCT amino acids may be fused with those of another protein, such as KLH, and antibodies to the chimeric molecule may be produced.

Monoclonal antibodies to HJNCT may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique. (See, e.g., Kohler, G. et al. (1975) Nature 256:495-497; Kozbor, D. et al. (1985) J. Immunol. Methods 81:31-42; Cote, R.J. et al. (1983) Proc. Natl. Acad. Sci. USA 80:2026-2030; Cole, S.P. et al. (1984) Mol. Cell Biol. 62:109-120.)

In addition, techniques developed for the production of "chimeric antibodies," such as the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used. (See, e.g., Morrison, S.L. et al. (1984) Proc. Natl. Acad. Sci. USA 81:6851-6855; Neuberger, M.S. et al. (1984) Nature 312:604-608; and Takeda, S. et al. (1985) Nature 314:452-454.) Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce HJNCT-specific single chain antibodies.

Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries. (See, e.g., Burton, D.R. (1991) Proc. Natl. Acad. Sci. USA 88:10134-10137.)

Antibodies may also be produced by inducing in vivo production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature. (See, e.g., Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. USA 86:3833-3837; Winter, G. et al. (1991) Nature 349:293-299.)

Antibody fragments which contain specific binding sites for HJNCT may also be generated. For example, such fragments include, but are not limited to, F(ab')2 fragments produced by pepsin digestion of the antibody molecule and Fab fragments generated by reducing the disulfide bridges of the F(ab')2 fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity. (See, e.g., Huse, W.D. et al. (1989) Science 246:1275-1281.)

Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between HJNCT and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering HJNCT epitopes is preferred, but a competitive binding assay may also be employed (Pound, supra).

Various methods such as Scatchard analysis in conjunction with radioimmunoassay techniques may be used to assess the affinity of antibodies for HJNCT. Affinity is expressed as an association constant, K_a, which is defined as the molar concentration of HJNCT-antibody complex divided by the molar concentrations of free antigen and free antibody under equilibrium conditions. The K_a determined for a preparation of polyclonal antibodies, which are heterogeneous in their affinities for multiple HJNCT epitopes, represents the average affinity, or avidity, of the antibodies for HJNCT. The K_a determined for a preparation of monoclonal antibodies, which are monospecific for a particular HJNCT epitope, represents a true measure of affinity. High-affinity antibody preparations with K_a ranging from about 10^9 to 10^{12} L/mole are preferred for use in immunoassays in which the HJNCT-antibody complex must withstand rigorous manipulations. Low-affinity antibody preparations with K_a ranging from about 10^6 to 10^7 L/mole are preferred for use in immunopurification and similar procedures which ultimately require dissociation of HJNCT, preferably in active form, from the antibody (Catty, D. (1988) Antibodies, Volume I: A Practical Approach, IRL Press, Washington, DC; Liddell, J.E. and A. Cryer (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY).

The titer and avidity of polyclonal antibody preparations may be further evaluated to determine the quality and suitability of such preparations for certain downstream applications. For example, a polyclonal antibody preparation containing at least 1-2 mg specific antibody/ml, preferably 5-10 mg specific antibody/ml, is preferred for use in procedures requiring precipitation of HJNCT-antibody complexes. Procedures for evaluating antibody specificity, titer, and avidity, and guidelines for antibody quality and usage in various applications, are generally available. (See, e.g., Catty, supra, and Coligan et al. supra.)

In another embodiment of the invention, the polynucleotides encoding HJNCT, or any fragment or complement thereof, may be used for therapeutic purposes. In one aspect, the complement of the polynucleotide encoding HJNCT may be used in situations in which it would be desirable to block the transcription of the mRNA. In particular, cells may be transformed with sequences complementary to polynucleotides encoding HJNCT. Thus, complementary molecules or fragments may be used to modulate HJNCT activity, or to achieve regulation of gene function. Such technology is now well known in the art, and sense or antisense oligonucleotides or larger fragments can be designed from various locations along the coding or control regions of sequences encoding HJNCT.

Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of nucleotide sequences to the targeted organ, tissue, or cell population. Methods which are well known to those skilled in the art can be used to construct vectors to express nucleic acid sequences complementary to the polynucleotides encoding HJNCT. (See, e.g., Sambrook, supra; Ausubel, 1995, supra.)

Genes encoding HJNCT can be turned off by transforming a cell or tissue with expression vectors which express high levels of a polynucleotide, or fragment thereof, encoding HJNCT. Such constructs may be used to introduce untranslatable sense or antisense sequences into a cell. Even in the absence of integration into the DNA, such vectors may continue to transcribe RNA molecules until they are disabled by endogenous nucleases. Transient expression may last for a month or more with a non-replicating vector, and may last even longer if appropriate replication elements are part of the vector system.

As mentioned above, modifications of gene expression can be obtained by designing complementary sequences or antisense molecules (DNA, RNA, or PNA) to the control, 5', or regulatory regions of the gene encoding HJNCT. Oligonucleotides derived from the transcription initiation site, e.g., between about positions -10 and +10 from the start site, are preferred. Similarly, inhibition can be achieved using triple helix base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature. (See, e.g., Gee, J.E. et al. (1994) in Huber, B.E. and B.I. Carr, Molecular and

Immunologic Approaches, Futura Publishing, Mt. Kisco NY, pp. 163-177.) A complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. For example, engineered hammerhead motif ribozyme molecules may specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding HJNCT.

Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites, including the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides, corresponding to the region of the target gene containing the cleavage site, may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Complementary ribonucleic acid molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA sequences encoding HJNCT. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize complementary RNA, constitutively or inducibly, can be introduced into cell lines, cells, or tissues.

RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule, or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

Many methods for introducing vectors into cells or tissues are available and equally suitable for use in vivo, in vitro, and ex vivo. For ex vivo therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection, by liposome injections, or by polycationic amino polymers may be achieved using methods which are well known in the art. (See, e.g., Goldman, C.K. et al. (1997) Nat. Biotechnol.

15:462-466.)

Any of the therapeutic methods described above may be applied to any subject in need of such therapy, including, for example, mammals such as dogs, cats, cows, horses, rabbits, monkeys, and most preferably, humans.

An additional embodiment of the invention relates to the administration of a pharmaceutical or sterile composition, in conjunction with a pharmaceutically acceptable carrier, for any of the therapeutic effects discussed above. Such pharmaceutical compositions may consist of HJNCT, antibodies to HJNCT, and mimetics, agonists, antagonists, or inhibitors of HJNCT. The compositions may be administered alone or in combination with at least one other agent, such as a stabilizing compound, which may be administered in any sterile, biocompatible pharmaceutical carrier including, but not limited to, saline, buffered saline, dextrose, and water. The compositions may be administered to a patient alone, or in combination with other agents, drugs, or hormones.

The pharmaceutical compositions utilized in this invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal means.

In addition to the active ingredients, these pharmaceutical compositions may contain suitable pharmaceutically-acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. Further details on techniques for formulation and administration may be found in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing, Easton PA).

Pharmaceutical compositions for oral administration can be formulated using pharmaceutically acceptable carriers well known in the art in dosages suitable for oral administration. Such carriers enable the pharmaceutical compositions to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions, and the like, for ingestion by the patient.

Pharmaceutical preparations for oral use can be obtained through combining active compounds with solid excipient and processing the resultant mixture of granules (optionally, after grinding) to obtain tablets or dragee cores. Suitable auxiliaries can be added, if desired. Suitable excipients include carbohydrate or protein fillers, such as sugars, including lactose, sucrose, mannitol, and sorbitol; starch from corn, wheat, rice, potato, or other plants; cellulose, such as methyl cellulose, hydroxypropylmethyl-cellulose, or sodium carboxymethylcellulose; gums, including arabic and tragacanth; and proteins, such as gelatin and collagen. If desired, disintegrating or solubilizing agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, and alginic acid or a salt thereof, such as sodium alginate.

Dragee cores may be used in conjunction with suitable coatings, such as concentrated sugar solutions, which may also contain gum arabic, talc, polyvinylpyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for product identification or to characterize the quantity of active compound, i.e., dosage.

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a coating, such as glycerol or sorbitol. Push-fit capsules can contain active ingredients mixed with fillers or binders, such as lactose or starches, lubricants, such as talc or magnesium stearate, and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid, or liquid polyethylene glycol with or without stabilizers.

Pharmaceutical formulations suitable for parenteral administration may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks' solution, Ringer's solution, or physiologically buffered saline. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils, such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate, triglycerides, or liposomes. Non-lipid polycationic amino polymers may also be used for delivery. Optionally, the suspension may also contain suitable stabilizers or agents to increase the solubility of the compounds and allow for the preparation of highly concentrated solutions.

For topical or nasal administration, penetrants appropriate to the particular barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

The pharmaceutical compositions of the present invention may be manufactured in a manner that is known in the art, e.g., by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping, or lyophilizing processes.

The pharmaceutical composition may be provided as a salt and can be formed with many acids, including but not limited to, hydrochloric, sulfuric, acetic, lactic, tartaric, malic, and succinic acids. Salts tend to be more soluble in aqueous or other protonic solvents than are the corresponding free base forms. In other cases, the preferred preparation may be a lyophilized powder which may contain any or all of the following: 1 mM to 50 mM histidine, 0.1% to 2% sucrose, and 2% to 7% mannitol, at a pH range of 4.5 to 5.5, that is combined with buffer prior to use.

After pharmaceutical compositions have been prepared, they can be placed in an appropriate container and labeled for treatment of an indicated condition. For administration of HJNCT, such labeling would include amount, frequency, and method of administration.

Pharmaceutical compositions suitable for use in the invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of neoplastic cells or in animal models such as mice, rats, rabbits, dogs, or pigs. An animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

A therapeutically effective dose refers to that amount of active ingredient, for example HJNCT or fragments thereof, antibodies of HJNCT, and agonists, antagonists or inhibitors of HJNCT, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the ED_{50} (the dose therapeutically effective in 50% of the population) or LD_{50} (the dose lethal to 50% of the population) statistics. The dose ratio of toxic to therapeutic effects is the therapeutic index, which can be expressed as the LD_{50}/ED_{50} ratio. Pharmaceutical compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the ED_{50} with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the subject requiring treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of the disease state, the general health of the subject, the age, weight, and gender of the subject, time and frequency of administration, drug combination(s), reaction sensitivities, and response to therapy. Long-acting pharmaceutical compositions may be administered every 3 to 4 days, every week, or biweekly depending on the half-life and clearance rate of the particular formulation.

Normal dosage amounts may vary from about 0.1 μg to 100,000 μg, up to a total dose of about 1 gram, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

DIAGNOSTICS

In another embodiment, antibodies which specifically bind HJNCT may be used for the diagnosis of disorders characterized by expression of HJNCT, or in assays to monitor patients being treated with HJNCT or agonists, antagonists, or inhibitors of HJNCT. Antibodies useful for diagnostic purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for HJNCT include methods which utilize the antibody and a label to detect HJNCT in human body fluids or in extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by covalent or non-covalent attachment of a reporter molecule. A wide variety of reporter molecules, several of which are described above, are known in the art and may be used.

A variety of protocols for measuring HJNCT, including ELISAs, RIAs, and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of HJNCT expression. Normal or standard values for HJNCT expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, preferably human, with antibody to HJNCT under conditions suitable for complex formation. The amount of standard complex formation may be quantitated by various methods, preferably by photometric means. Quantities of HJNCT expressed in subject, control, and disease samples from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

In another embodiment of the invention, the polynucleotides encoding HJNCT may be used for diagnostic purposes. The polynucleotides which may be used include oligonucleotide sequences, complementary RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantitate gene expression in biopsied tissues in which expression of HJNCT may be correlated with disease. The diagnostic assay may be used to determine absence, presence, and excess expression of HJNCT, and to monitor regulation of HJNCT levels during therapeutic intervention.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding HJNCT or closely related molecules may be used to identify nucleic acid sequences which encode HJNCT. The specificity of the probe, whether it is made from a highly specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a conserved motif, and the stringency of the hybridization or amplification (maximal, high, intermediate, or low), will determine whether the probe identifies only naturally occurring sequences encoding HJNCT, allelic variants, or related sequences.

Probes may also be used for the detection of related sequences, and should preferably have at least 50% sequence identity to any of the HJNCT encoding sequences. The hybridization probes of the subject invention may be DNA or RNA and may be derived from the sequence of SEQ ID NO:5-8 or from genomic sequences including promoters, enhancers, and introns of the HJNCT gene.

Means for producing specific hybridization probes for DNAs encoding HJNCT include the cloning of polynucleotide sequences encoding HJNCT or HJNCT derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by means of the addition of the appropriate RNA polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a variety of reporter groups, for example, by radionuclides such as ^{32}P or ^{35}S, or by enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Polynucleotide sequences encoding HJNCT may be used for the diagnosis of disorders associated with expression of HJNCT. Examples of such disorders include, but are not limited to, a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancer, including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus; an autoimmune/inflammatory disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyenodocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru,

Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis, inherited, metabolic, endocrine, and toxic myopathies, myasthenia gravis, periodic paralysis, mental disorders including mood, anxiety, and schizophrenic disorders, akathesia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, and Tourette's disorder; a developmental disorder such as renal tubular acidosis, anemia, Cushing's syndrome, achondroplastic dwarfism, Duchenne and Becker muscular dystrophy, epilepsy, gonadal dysgenesis, WAGR syndrome (Wilms' tumor, aniridia, genitourinary abnormalities, and mental retardation), Smith-Magenis syndrome, myelodysplastic syndrome, hereditary mucoepithelial dysplasia, hereditary keratodermas, hereditary neuropathies such as Charcot-Marie-Tooth disease and neurofibromatosis, hypothyroidism, hydrocephalus, seizure disorders such as Syndenham's chorea and cerebral palsy, spina bifida, Williams syndrome, anencephaly, craniorachischisis, congenital glaucoma, cataract, sensorineural hearing loss, and any disorder associated with cell growth and differentiation, embryogenesis, and morphogenesis involving any tissue, organ, or system of a subject, e.g., the brain, adrenal gland, kidney, skeletal or reproductive system; a vesicle trafficking disorder such as cystic fibrosis, glucose-galactose malabsorption syndrome, hypercholesterolemia, diabetes mellitus, diabetes insipidus, hyper- and hypoglycemia, Grave's disease, goiter, Cushing's disease, and Addison's disease, other conditions associated with abnormal vesicle trafficking, including acquired immunodeficiency syndrome (AIDS), allergies including hay fever, asthma, and urticaria (hives); autoimmune hemolytic anemia, proliferative glomerulonephritis, inflammatory bowel disease, multiple sclerosis, myasthenia gravis, rheumatoid and osteoarthritis; scleroderma, Chediak-Higashi and Sjogren's syndromes; systemic lupus erythematosus, toxic shock syndrome, traumatic tissue damage, and viral, bacterial, fungal, helminthic, and protozoal infections; a reproductive disorder such as a disorder of prolactin production; infertility, including tubal disease, ovulatory defects, and endometriosis; disruptions of the estrous cycle, disruptions of the menstrual cycle, polycystic ovary syndrome, ovarian hyperstimulation syndrome, endometrial and ovarian tumors, uterine fibroids, autoimmune disorders, ectopic pregnancies, and teratogenesis; cancer of the breast, fibrocystic breast disease, and galactorrhea; disruptions of spermatogenesis, abnormal sperm physiology, cancer of the testis, cancer of the prostate, benign prostatic hyperplasia, prostatitis, Peyronie's disease, impotence, carcinoma of the male breast, and gynecomastia; a gastrointestinal disorder such as dysphagia, peptic esophagitis, esophageal spasm, esophageal stricture, esophageal carcinoma, dyspepsia,

indigestion, gastritis, gastric carcinoma, anorexia, nausea, emesis, gastroparesis, antral or pyloric edema, abdominal angina, pyrosis, gastroenteritis, intestinal obstruction, infections of the intestinal tract, peptic ulcer, gastric ulcer, duodenal ulcer, cholelithiasis, cholecystitis, cholestasis, pancreatitis, pancreatic carcinoma, biliary tract disease, hepatitis, hyperbilirubinemia, cirrhosis, passive congestion of the liver, hepatoma, infectious colitis, ulcerative colitis, ulcerative proctitis, Crohn's disease, Whipple's disease, Mallory-Weiss syndrome, colonic carcinoma, colonic obstruction, irritable bowel syndrome, short bowel syndrome, diarrhea, constipation, gastrointestinal hemorrhage, acquired immunodeficiency syndrome (AIDS) enteropathy, jaundice, hepatic encephalopathy, hepatorenal syndrome, hepatic steatosis, hemochromatosis, Wilson's disease, $alpha_1$-antitrypsin deficiency, Reye's syndrome, primary sclerosing cholangitis, liver infarction, portal vein obstruction and thrombosis, centrilobular necrosis, peliosis hepatis, hepatic vein thrombosis, veno-occlusive disease, preeclampsia, eclampsia, acute fatty liver of pregnancy, intrahepatic cholestasis of pregnancy, and hepatic tumors including nodular hyperplasias, adenomas, and carcinomas; and a renal disorder such as renal amyloidosis, hypertension, primary aldosteronism, Addison's disease, renal failure, glomerulonephritis, chronic glomerulonephritis, tubulointerstitial nephritis, cystic disorders of the kidney and dysplastic malformations such as polycystic disease, renal dysplasias, and cortical or medullary cysts, inherited polycystic renal diseases (PRD) such as recessive and autosomal dominant PRD, medullary cystic disease, medullary sponge kidney and tubular dysplasia, Alport's syndrome, non-renal cancers which affect renal physiology such as bronchogenic tumors of the lungs or tumors of the basal region of the brain, multiple myeloma, adenocarcinomas of the kidney, metastatic renal carcinoma, and nephrotoxic disorders including any functional or morphologic change in the kidney produced by any pharmaceutical, chemical, or biological agent that is ingested, injected, inhaled, or absorbed. Some broad categories of common nephrotoxic agents are heavy metals, all classes of antibiotics, analgesics, solvents, oxalosis-inducing agents, anticancer drugs, herbicides and pesticides, botanicals and biologicals, and antiepileptics. The polynucleotide sequences encoding HJNCT may be used in Southern or northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; in dipstick, pin, and multiformat ELISA-like assays; and in microarrays utilizing fluids or tissues from patients to detect altered HJNCT expression. Such qualitative or quantitative methods are well known in the art.

In a particular aspect, the nucleotide sequences encoding HJNCT may be useful in assays that detect the presence of associated disorders, particularly those mentioned above. The nucleotide sequences encoding HJNCT may be labeled by standard methods and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantitated and compared with a standard value. If the amount of signal in the patient sample is significantly altered in comparison to a control sample then

the presence of altered levels of nucleotide sequences encoding HJNCT in the sample indicates the presence of the associated disorder. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or to monitor the treatment of an individual patient.

In order to provide a basis for the diagnosis of a disorder associated with expression of HJNCT, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence, or a fragment thereof, encoding HJNCT, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with values from an experiment in which a known amount of a substantially purified polynucleotide is used. Standard values obtained in this manner may be compared with values obtained from samples from patients who are symptomatic for a disorder. Deviation from standard values is used to establish the presence of a disorder.

Once the presence of a disorder is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to determine if the level of expression in the patient begins to approximate that which is observed in the normal subject. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

With respect to cancer, the presence of an abnormal amount of transcript (either under- or overexpressed) in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the cancer.

Additional diagnostic uses for oligonucleotides designed from the sequences encoding HJNCT may involve the use of PCR. These oligomers may be chemically synthesized, generated enzymatically, or produced in vitro. Oligomers will preferably contain a fragment of a polynucleotide encoding HJNCT, or a fragment of a polynucleotide complementary to the polynucleotide encoding HJNCT, and will be employed under optimized conditions for identification of a specific gene or condition. Oligomers may also be employed under less stringent conditions for detection or quantitation of closely related DNA or RNA sequences.

Methods which may also be used to quantify the expression of HJNCT include radiolabeling or biotinylating nucleotides, coamplification of a control nucleic acid, and interpolating results from standard curves. (See, e.g., Melby, P.C. et al. (1993) J. Immunol. Methods 159:235-244; Duplaa, C. et al. (1993) Anal. Biochem. 212:229-236.) The speed of quantitation of multiple samples may be accelerated by

running the assay in an ELISA format where the oligomer of interest is presented in various dilutions and a spectrophotometric or colorimetric response gives rapid quantitation.

In further embodiments, oligonucleotides or longer fragments derived from any of the polynucleotide sequences described herein may be used as targets in a microarray. The microarray can be used to monitor the expression level of large numbers of genes simultaneously and to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a disorder, to diagnose a disorder, and to develop and monitor the activities of therapeutic agents.

Microarrays may be prepared, used, and analyzed using methods known in the art. (See, e.g., Brennan, T.M. et al. (1995) U.S. Patent No. 5,474,796; Schena, M. et al. (1996) Proc. Natl. Acad. Sci. USA 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/251116; Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R.A. et al. (1997) Proc. Natl. Acad. Sci. USA 94:2150-2155; and Heller, M.J. et al. (1997) U.S. Patent No. 5,605,662.)

In another embodiment of the invention, nucleic acid sequences encoding HJNCT may be used to generate hybridization probes useful in mapping the naturally occurring genomic sequence. The sequences may be mapped to a particular chromosome, to a specific region of a chromosome, or to artificial chromosome constructions, e.g., human artificial chromosomes (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1 constructions, or single chromosome cDNA libraries. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355; Price, C.M. (1993) Blood Rev. 7:127-134; and Trask, B.J. (1991) Trends Genet. 7:149-154.)

Fluorescent in situ hybridization (FISH) may be correlated with other physical chromosome mapping techniques and genetic map data. (See, e.g., Heinz-Ulrich, et al. (1995) in Meyers, supra, pp. 965-968.) Examples of genetic map data can be found in various scientific journals or at the Online Mendelian Inheritance in Man (OMIM) site. Correlation between the location of the gene encoding HJNCT on a physical chromosomal map and a specific disorder, or a predisposition to a specific disorder, may help define the region of DNA associated with that disorder. The nucleotide sequences of the invention may be used to detect differences in gene sequences among normal, carrier, and affected individuals.

In situ hybridization of chromosomal preparations and physical mapping techniques, such as linkage analysis using established chromosomal markers, may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the number or arm of a particular human chromosome is not known. New sequences can be assigned to chromosomal arms by physical mapping. This provides valuable information to investigators searching for disease genes using positional cloning or other gene discovery

techniques. Once the disease or syndrome has been crudely localized by genetic linkage to a particular genomic region, e.g., ataxia-telangiectasia to 11q22-23, any sequences mapping to that area may represent associated or regulatory genes for further investigation. (See, e.g., Gatti, R.A. et al. (1988) Nature 336:577-580.) The nucleotide sequence of the subject invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc., among normal, carrier, or affected individuals.

In another embodiment of the invention, HJNCT, its catalytic or immunogenic fragments, or oligopeptides thereof can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes between HJNCT and the agent being tested may be measured.

Another technique for drug screening provides for high throughput screening of compounds having suitable binding affinity to the protein of interest. (See, e.g., Geysen, et al. (1984) PCT application WO84/03564.) In this method, large numbers of different small test compounds are synthesized on a solid substrate. The test compounds are reacted with HJNCT, or fragments thereof, and washed. Bound HJNCT is then detected by methods well known in the art. Purified HJNCT can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding HJNCT specifically compete with a test compound for binding HJNCT. In this manner, antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with HJNCT.

In additional embodiments, the nucleotide sequences which encode HJNCT may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

Without further elaboration, it is believed that one skilled in the art can, using the preceding description, utilize the present invention to its fullest extent. The following preferred specific embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

The disclosures of all patents, applications, and publications mentioned above and below, in particular U.S. Ser. No. [Attorney Docket No. PF-0590 P, filed September 25, 1998], U.S. Ser No. [Attorney Docket No. PF-0613 P, filed October 13, 1998], and U.S. Ser. No. [Attorney Docket No. PF-

0700 P, filed May 4, 1999], are hereby expressly incorporated by reference.

EXAMPLES

I. Construction of cDNA Libraries

RNA was isolated from tissues described in Table 4. For construction of the LUNGNOT23 and OVARTUT02 cDNA libraries, the frozen tissue was homogenized and lysed in TRIZOL reagent (1gm tissue/10 ml TRIZOL reagent; Life Technologies), a monophasic solution of phenol and guanidine isothiocyanate, using a Polytron PT-3000 homogenizer (Brinkman Instruments, Westbury NY). After a brief incubation on ice, chloroform was added (1:5 v/v), and the mixture was centrifuged to separate the phases. The upper aqueous phase was removed to a fresh tube, and isopropanol was added to precipitate RNA. The RNA was resuspended in RNase-free water and treated with DNase. The RNA was re-extracted as necessary with acid phenol-chloroform to increase purity, and the RNA was reprecipitated with sodium acetate and ethanol. For construction of the BLADNOT04 cDNA library, the frozen tissue was homogenized and lysed using a Polytron PT-3000 homogenizer (Brinkmann Instruments) in guanidinium isothiocyanate solution. The lysate was centrifuged over a 5.7 M CsCl cushion using an SW28 rotor in an L8-70M ultracentrifuge (Beckman Instruments, Fullerton CA) for 18 hours at 25,000 rpm at ambient temperature. RNA was extracted with acid phenol pH 4.7, precipitated using 0.3 M sodium acetate and 2.5 volumes of ethanol, resuspended in RNAse-free water, and treated with DNase at 37°C. RNA extraction and precipitation steps were repeated. For construction of tbe OVARDIT04 cDNA library, RNA was isolated directly from tissue lysates using other RNA isolation kits, e.g., the POLY(A)PURE mRNA purification kit (Ambion, Austin TX), or HPLC.

From each RNA preparation, poly(A+) RNA was isolated using the OLIGOTEX kit (QIAGEN, Chatsworth CA). Poly(A+) RNA was used for cDNA synthesis and construction of each cDNA library according to the recommended protocols in the SUPERSCRIPT plasmid system (Life Technologies). The cDNAs were fractionated on a SEPHAROSE CL4B column (Amersham Pharmacia Biotech), and those cDNAs exceeding 400 bp were ligated into pINCY (Incyte Pharmaceuticals, Palo Alto CA). Recombinant plasmids were transformed into DH5α competent cells or ElectroMAX®cells (Life Technologies).

II. Isolation of cDNA Clones

Plasmid DNA was released from the cells and purified using the R.E.A.L. PREP 96 plasmid purification kit (QIAGEN). The recommended protocol was employed except for the following changes: 1) the bacteria were cultured in 1 ml of sterile Terrific Broth (Life Technologies) with carbenicillin at 25 mg/l and glycerol at 0.4%; 2) after the cultures were incubated for 19 hours, the cells were lysed with 0.3 ml of lysis buffer; and 3) following isopropanol precipitation, the plasmid DNA pellets were each

resuspended in 0.1 ml of distilled water. The DNA samples were stored at 4°C.

III. Sequencing and Analysis

cDNA sequencing reactions were processed using standard methods or high-throughput instrumentation such as the ABI CATALYST 800 (Perkin-Elmer) thermal cycler or the PTC-200 thermal cycler (MJ Research) in conjunction with the HYDRA microdispenser (Robbins Scientific) or the MICROLAB 2200 (Hamilton) liquid transfer system. cDNA sequencing reactions were prepared using reagents provided by Amersham Pharmacia Biotech or supplied in ABI sequencing kits such as the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Perkin-Elmer). Electrophoretic separation of cDNA sequencing reactions and detection of labeled polynucleotides were carried out using the MEGABACE 1000 DNA sequencing system (Molecular Dynamics); the ABI PRISM 373 or 377 sequencing system (Perkin-Elmer) in conjunction with standard ABI protocols and base calling software; or other sequence analysis systems known in the art. Reading frames within the cDNA sequences were identified using standard methods (reviewed in Ausubel, 1997, supra, unit 7.7). Some of the cDNA sequences were selected for extension using the techniques disclosed in Example V.

The polynucleotide sequences derived from cDNA sequencing were assembled and analyzed using a combination of software programs which utilize algorithms well known to those skilled in the art. Table 5 summarizes the tools, programs, and algorithms used and provides applicable descriptions, references, and threshold parameters. The first column of Table 5 shows the tools, programs, and algorithms used, the second column provides brief descriptions thereof, the third column presents appropriate references, all of which are incorporated by reference herein in their entirety, and the fourth column presents, where applicable, the scores, probability values, and other parameters used to evaluate the strength of a match between two sequences (the higher the score, the greater the homology between two sequences). Sequences were analyzed using MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA) and LASERGENE software (DNASTAR). Polynucleotide and polypeptide sequence alignments were generated using the default parameters specified by the clustal algorithm as incorporated into the MEGALIGN multisequence alignment program (DNASTAR), which also calculates the percent identity between aligned sequences.

The polynucleotide sequences were validated by removing vector, linker, and polyA sequences and by masking ambiguous bases, using algorithms and programs based on BLAST, dynamic programing, and dinucleotide nearest neighbor analysis. The sequences were then queried against a selection of public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases, and BLOCKS to acquire annotation using programs based on BLAST, FASTA, and BLIMPS. The sequences were assembled into full length polynucleotide sequences using programs based on Phred, Phrap, and Consed, and were screened for open reading frames using programs based on GeneMark,

BLAST, and FASTA. The full length polynucleotide sequences were translated to derive the corresponding full length amino acid sequences, and these full length sequences were subsequently analyzed by querying against databases such as the GenBank databases (described above), SwissProt, BLOCKS, PRINTS, Prosite, and Hidden Markov Model (HMM)-based protein family databases such as PFAM. HMM is a probabilistic approach which analyzes consensus primary structures of gene families. (See, e.g., Eddy, S.R. (1996) Curr. Opin. Struct. Biol. 6:361-365.)

The programs described above for the assembly and analysis of full length polynucleotide and amino acid sequences were also used to identify polynucleotide sequence fragments from SEQ ID NO:5-8. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies were described in The Invention section above.

IV. Northern Analysis

Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs from a particular cell type or tissue have been bound. (See, e.g., Sambrook, supra, ch. 7; Ausubel, 1995, supra, ch. 4 and 16.)

Analogous computer techniques applying BLAST were used to search for identical or related molecules in nucleotide databases such as GenBank or LIFESEQ (Incyte Pharmaceuticals). This analysis is much faster than multiple membrane-based hybridizations. In addition, the sensitivity of the computer search can be modified to determine whether any particular match is categorized as exact or similar. The basis of the search is the product score, which is defined as:

$$\frac{\%\text{ sequence identity} \times \%\text{ maximum BLAST score}}{100}$$

The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. For example, with a product score of 40, the match will be exact within a 1% to 2% error, and, with a product score of 70, the match will be exact. Similar molecules are usually identified by selecting those which show product scores between 15 and 40, although lower scores may identify related molecules.

The results of northern analyses are reported as a percentage distribution of libraries in which the transcript encoding HJNCT occurred. Analysis involved the categorization of cDNA libraries by organ/tissue and disease. The organ/tissue categories included cardiovascular, dermatologic, developmental, endocrine, gastrointestinal, hematopoietic/immune, musculoskeletal, nervous, reproductive, and urologic. The disease/condition categories included cancer, inflammation/trauma, cell proliferation, neurological, and pooled. For each category, the number of libraries expressing the sequence of interest was counted and divided by the total number of libraries across all categories.

Percentage values of tissue-specific and disease- or condition-specific expression are reported in Table 3.

V. Extension of HJNCT Encoding Polynucleotides

The full length nucleic acid sequence of SEQ ID NO:5 was produced by extension of an appropriate fragment of the full length molecule, using oligonucleotide primers designed from this fragment. One primer was synthesized to initiate extension of an antisense polynucleotide, and the other was synthesized to initiate extension of a sense polynucleotide. Primers were used to facilitate the extension of the known sequence "outward" generating amplicons containing new unknown nucleotide sequence for the region of interest. The initial primers were designed from the cDNA using OLIGO 4.06 software (National Biosciences, Plymouth, MN), or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the target sequence at temperatures of about 68°C to about 72°C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations was avoided.

Selected human cDNA libraries (Life Technologies) were used to extend the sequence. If more than one extension is necessary or desired, additional sets of primers are designed to further extend the known region.

High fidelity amplification was obtained by following the instructions for the XL-PCR kit (Perkin-Elmer Corp., Norwalk, CT) and thoroughly mixing the enzyme and reaction mix. PCR was performed using the PTC-200 thermal cycler (MJ Research, Inc., Watertown, MA), beginning with 40 pmol of each primer and the recommended concentrations of all other components of the kit, with the following parameters:

Step 1	94° C for 1 min (initial denaturation)
Step 2	65° C for 1 min
Step 3	68° C for 6 min
Step 4	94° C for 15 sec
Step 5	65° C for 1 min
Step 6	68° C for 7 min
Step 7	Repeat steps 4 through 6 for an additional 15 cycles
Step 8	94° C for 15 sec
Step 9	65° C for 1 min
Step 10	68° C for 7:15 min
Step 11	Repeat steps 8 through 10 for an additional 12 cycles
Step 12	72° C for 8 min
Step 13	4° C (and holding)

A 5 μl to 10 μl aliquot of the reaction mixture was analyzed by electrophoresis on a low concentration (about 0.6% to 0.8%) agarose mini-gel to determine which reactions were successful in extending the sequence. Bands thought to contain the largest products were excised from the gel, purified using QIAQUICK DNA gel purification kit (Qiagen, Inc.), and trimmed of overhangs using Klenow

enzyme to facilitate religation and cloning.

After ethanol precipitation, the products were redissolved in 13 μl of ligation buffer, 1μl T4-DNA ligase (15 units) and 1μl T4 polynucleotide kinase were added, and the mixture was incubated at room temperature for 2 to 3 hours, or overnight at 16° C. Competent E. coli cells (in 40 μl of appropriate media) were transformed with 3 μl of ligation mixture and cultured in 80 μl of SOC medium. (See, e.g., Sambrook, supra, Appendix A, p. 2.) After incubation for one hour at 37°C, the E. coli mixture was plated on Luria Bertani (LB) agar (See, e.g., Sambrook, supra, Appendix A, p. 1) containing carbenicillin (2x carb). The following day, several colonies were randomly picked from each plate and cultured in 150 μl of liquid LB/2x carb medium placed in an individual well of an appropriate commercially-available sterile 96-well microtiter plate. The following day, 5 μl of each overnight culture was transferred into a non-sterile 96-well plate and, after dilution 1:10 with water, 5 μl from each sample was transferred into a PCR array.

For PCR amplification, 18 μl of concentrated PCR reaction mix (3.3x) containing 4 units of rTth DNA polymerase, a vector primer, and one or both of the gene specific primers used for the extension reaction were added to each well. Amplification was performed using the following conditions:

Step 1	94° C for 60 sec
Step 2	94° C for 20 sec
Step 3	55° C for 30 sec
Step 4	72° C for 90 sec
Step 5	Repeat steps 2 through 4 for an additional 29 cycles
Step 6	72° C for 180 sec
Step 7	4° C (and holding)

Aliquots of the PCR reactions were run on agarose gels together with molecular weight markers. The sizes of the PCR products were compared to the original partial cDNAs, and appropriate clones were selected, ligated into plasmid, and sequenced.

In like manner, the nucleotide sequence of SEQ ID NO:5 is used to obtain 5' regulatory sequences using the procedure above, oligonucleotides designed for 5' extension, and an appropriate genomic library.

The full length nucleic acid sequences of SEQ ID NO:6-8 were produced by extension of appropriate fragments of the full length molecules using oligonucleotide primers designed from these fragments. Primers were synthesized to initiate either 5' extension of the known fragments or 3' extension of the known fragments. The initial primers were designed using OLIGO 4.06 software (National Biosciences), or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the target sequence at temperatures of about 68°C to about 72°C. Any stretch of nucleotides which would result in hairpin structures and primer-primer

dimerizations was avoided.

Selected human cDNA libraries were used to extend the sequence. If more than one extension was necessary or desired, additional or nested sets of primers were designed.

High fidelity amplification was obtained by PCR using methods well known in the art. PCR was performed in 96-well plates using the PTC-200 thermal cycler (MJ Research, Inc.). The reaction mix contained DNA template, 200 nmol of each primer, reaction buffer containing Mg^{2+}, $(NH_4)_2SO_4$, and β-mercaptoethanol, Taq DNA polymerase (Amersham Pharmacia Biotech), ELONGASE enzyme (Life Technologies), and Pfu DNA polymerase (Stratagene), with the following parameters for primer pair PCI A and PCI B: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C. In the alternative, the parameters for primer pair T7 and SK+ were as follows: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 57°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C.

The concentration of DNA in each well was determined by dispensing 100 μl PICOGREEN quantitation reagent (0.25% (v/v) PICOGREEN; Molecular Probes, Eugene OR) dissolved in 1X TE and 0.5 μl of undiluted PCR product into each well of an opaque fluorimeter plate (Corning Costar, Acton MA), allowing the DNA to bind to the reagent. The plate was scanned in a Fluoroskan II (Labsystems Oy, Helsinki, Finland) to measure the fluorescence of the sample and to quantify the concentration of DNA. A 5 μl to 10 μl aliquot of the reaction mixture was analyzed by electrophoresis on a 1 % agarose mini-gel to determine which reactions were successful in extending the sequence.

The extended nucleotides were desalted and concentrated, transferred to 384-well plates, digested with CviJI cholera virus endonuclease (Molecular Biology Research, Madison WI), and sonicated or sheared prior to religation into pUC 18 vector (Amersham Pharmacia Biotech). For shotgun sequencing, the digested nucleotides were separated on low concentration (0.6 to 0.8%) agarose gels, fragments were excised, and agar digested with Agar ACE (Promega). Extended clones were religated using T4 ligase (New England Biolabs, Beverly MA) into pUC 18 vector (Amersham Pharmacia Biotech), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction site overhangs, and transfected into competent E. coli cells. Transformed cells were selected on antibiotic-containing media, individual colonies were picked and cultured overnight at 37°C in 384-well plates in LB/2x carb liquid media.

The cells were lysed, and DNA was amplified by PCR using Taq DNA polymerase (Amersham Pharmacia Biotech) and Pfu DNA polymerase (Stratagene) with the following parameters: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 72°C, 2 min; Step 5: steps 2, 3, and 4 repeated 29 times; Step 6: 72°C, 5 min; Step 7: storage at 4°C. DNA was quantified by PICOGREEN reagent

(Molecular Probes) as described above. Samples with low DNA recoveries were reamplified using the same conditions as described above. Samples were diluted with 20% dimethysulphoxide (1:2, v/v), and sequenced using DYENAMIC energy transfer sequencing primers and the DYENAMIC DIRECT kit (Amersham Pharmacia Biotech) or the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Perkin-Elmer).

In like manner, the nucleotide sequences of SEQ ID NO:6-8 are used to obtain 5' regulatory sequences using the procedure above, oligonucleotides designed for such extension, and an appropriate genomic library.

VI. Labeling and Use of Individual Hybridization Probes

Hybridization probes derived from SEQ ID NO:5-8 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base pairs, is specifically described, essentially the same procedure is used with larger nucleotide fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250 μCi of [γ-^{32}P] adenosine triphosphate (Amersham Pharmacia Biotech), and T4 polynucleotide kinase (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Pharmacia Biotech). An aliquot containing 10^7 counts per minute of the labeled probe is used in a typical membrane-based hybridization analysis of human genomic DNA digested with one of the following endonucleases: Ase I, Bgl II, Eco RI, Pst I, Xba1, or Pvu II (DuPont NEN).

The DNA from each digest is fractionated on a 0.7% agarose gel and transferred to nylon membranes (Nytran Plus, Schleicher & Schuell, Durham NH). Hybridization is carried out for 16 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature under increasingly stringent conditions up to 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate. Hybridization patterns are visualized using autoradiography and compared.

VII. Microarrays

A chemical coupling procedure and an ink jet device can be used to synthesize array elements on the surface of a substrate. (See, e.g., Baldeschweiler, supra.) An array analogous to a dot or slot blot may also be used to arrange and link elements to the surface of a substrate using thermal, UV, chemical, or mechanical bonding procedures. A typical array may be produced by hand or using available methods and machines and contain any appropriate number of elements. After hybridization, nonhybridized probes are removed and a scanner used to determine the levels and patterns of fluorescence. The degree of complementarity and the relative abundance of each probe which hybridizes to an element on the microarray may be assessed through analysis of the scanned images.

Full-length cDNAs, Expressed Sequence Tags (ESTs), or fragments thereof may comprise the elements of the microarray. Fragments suitable for hybridization can be selected using software well known in the art such as LASERGENE software (DNASTAR). Full-length cDNAs, ESTs, or fragments thereof corresponding to one of the nucleotide sequences of the present invention, or selected at random from a cDNA library relevant to the present invention, are arranged on an appropriate substrate, e.g., a glass slide. The cDNA is fixed to the slide using, e.g., UV cross-linking followed by thermal and chemical treatments and subsequent drying. (See, e.g., Schena, M. et al. (1995) Science 270:467-470; Shalon, D. et al. (1996) Genome Res. 6:639-645.) Fluorescent probes are prepared and used for hybridization to the elements on the substrate. The substrate is analyzed by procedures described above.

VIII. Complementary Polynucleotides

Sequences complementary to the HJNCT-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring HJNCT. Although use of oligonucleotides comprising from about 15 to 30 base pairs is described, essentially the same procedure is used with smaller or with larger sequence fragments. Appropriate oligonucleotides are designed using OLIGO 4.06 software (National Biosciences) and the coding sequence of HJNCT. To inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent promoter binding to the coding sequence. To inhibit translation, a complementary oligonucleotide is designed to prevent ribosomal binding to the HJNCT-encoding transcript.

IX. Expression of HJNCT

Expression and purification of HJNCT is achieved using bacterial or virus-based expression systems. For expression of HJNCT in bacteria, cDNA is subcloned into an appropriate vector containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA transcription. Examples of such promoters include, but are not limited to, the *trp-lac* (*tac*) hybrid promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3). Antibiotic resistant bacteria express HJNCT upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of HJNCT in eukaryotic cells is achieved by infecting insect or mammalian cell lines with recombinant Autographica californica nuclear polyhedrosis virus (AcMNPV), commonly known as baculovirus. The nonessential polyhedrin gene of baculovirus is replaced with cDNA encoding HJNCT by either homologous recombination or bacterial-mediated transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong polyhedrin promoter drives high levels of cDNA transcription. Recombinant baculovirus is used to infect Spodoptera frugiperda (Sf9) insect cells in most cases, or human hepatocytes, in some cases. Infection of the latter requires additional genetic modifications to baculovirus. (See Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-

3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945.)

In most expression systems, HJNCT is synthesized as a fusion protein with, e.g., glutathione S-transferase (GST) or a peptide epitope tag, such as FLAG or 6-His, permitting rapid, single-step, affinity-based purification of recombinant fusion protein from crude cell lysates. GST, a 26-kilodalton enzyme from Schistosoma japonicum, enables the purification of fusion proteins on immobilized glutathione under conditions that maintain protein activity and antigenicity (Amersham Pharmacia Biotech). Following purification, the GST moiety can be proteolytically cleaved from HJNCT at specifically engineered sites. FLAG, an 8-amino acid peptide, enables immunoaffinity purification using commercially available monoclonal and polyclonal anti-FLAG antibodies (Eastman Kodak). 6-His, a stretch of six consecutive histidine residues, enables purification on metal-chelate resins (QIAGEN). Methods for protein expression and purification are discussed in Ausubel (1995, supra, ch 10 and 16). Purified HJNCT obtained by these methods can be used directly in the following activity assay.

X. Demonstration of HJNCT Activity

An assay for HJNCT-1 activity measures the PDZ-induced clustering of transmembrane receptors (Ponting, supra). Cultured cell lines are cotransfected with cDNA encoding HJNCT-1 and either EGF receptor or NMDA receptor. Control cell lines are transfected with only one of the above cDNAs. Clustering of EGF receptors or NMDA receptors in the cotransfected cell lines is detected and quantified using commercially available antibody specific to these receptors in conjunction with indirect immunofluorescence and image analysis systems. The amount of receptor clustering is directly proportional to the amount of HJNCT-1 activity.

HJNCT-2 activity is demonstrated by its ability to bind GM130 in an in vitro assay (Barr (1998) supra). Coupled in vitro transcription-translation reactions are carried out with plasmids encoding HJNCT-2 and GM130 in the presence of radiolabeled [^{35}S]-methionine. Immunoprecipitations are performed on the in vitro transcription-translation reaction with antibodies to HJNCT-2. The immunoprecipitated material is analyzed by sodium dodecyl sulfate polyacrylamide gel electrophoresis and autoradiography to identify both HJNCT-2 and GM130 in the immunoprecipitates. The GM130 spot is cut out and counted in a radioisotope counter. The amount of radioactivity recovered is proportional to the activity of HJNCT-2 in the sample.

An assay for HJNCT-3 activity measures the ability of HJNCT-3 to induce the formation of tight junction strands when expressed in mouse fibroblasts (Furuse, M. et al. (1998) J. Cell Biol. 143:391-401). cDNA encoding HJNCT-3 is subcloned into a mammalian expression vector and transfected into mouse L fibroblasts. The transfected cells are analyzed by freeze-fracture electron microscopy after glutaraldehyde fixation and compared to untransfected cells. The presence of networks of strands and grooves at cell contact sites in transfected as compared to control cells is indicative of HJNCT-3 activity.

An assay for HJNCT-4 activity measures the disruption of cytoskeletal filament networks upon overexpression of HJNCT-4 in cultured cell lines (Rezniczek, G.A. et al. (1998) J. Cell Biol. 141:209-225). cDNA encoding HJNCT-4 is subcloned into a mammalian expression vector that drives high levels of cDNA expression. This construct is transfected into cultured cells, such as rat kangaroo PtK2 or rat bladder carcinoma 804G cells. Actin filaments and intermediate filaments such as keratin and vimentin are visualized by immunofluorescence microscopy using antibodies and techniques well known in the art. The configuration and abundance of cyoskeletal filaments can be assessed and quantified using confocal imaging techniques. In particular, the appearance of cell surface projections is indicative of HJNCT-4 activity.

XI. Functional Assays

HJNCT function is assessed by expressing the sequences encoding HJNCT at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression vector containing a strong promoter that drives high levels of cDNA expression. Vectors of choice include pCMV SPORT (Life Technologies) and pCR3.1 (Invitrogen, Carlsbad CA), both of which contain the cytomegalovirus promoter. 5-10 μg of recombinant vector are transiently transfected into a human cell line, preferably of endothelial or hematopoietic origin, using either liposome formulations or electroporation. 1-2 μg of an additional plasmid containing sequences encoding a marker protein are co-transfected. Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector. Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP; Clontech), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated, laser optics-based technique, is used to identify transfected cells expressing GFP or CD64-GFP and to evaluate the apoptotic state of the cells and other cellular properties. FCM detects and quantifies the uptake of fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies; and alterations in plasma membrane composition as measured by the binding of fluorescein-conjugated Annexin V protein to the cell surface. Methods in flow cytometry are discussed in Ormerod, M.G. (1994) Flow Cytometry, Oxford, New York NY.

The influence of HJNCT on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding HJNCT and either CD64 or CD64-GFP. CD64 and CD64-GFP are expressed on the surface of transfected cells and bind to conserved regions of human immunoglobulin G (IgG). Transfected cells are efficiently separated from nontransfected cells using

magnetic beads coated with either human IgG or antibody against CD64 (DYNAL, Lake Success NY). mRNA can be purified from the cells using methods well known by those of skill in the art. Expression of mRNA encoding HJNCT and other genes of interest can be analyzed by northern analysis or microarray techniques.

XII. Production of HJNCT Specific Antibodies

HJNCT substantially purified using polyacrylamide gel electrophoresis (PAGE; see, e.g., Harrington, M.G. (1990) Methods Enzymol. 182:488-495), or other purification techniques, is used to immunize rabbits and to produce antibodies using standard protocols.

Alternatively, the HJNCT amino acid sequence is analyzed using LASERGENE software (DNASTAR) to determine regions of high immunogenicity, and a corresponding oligopeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Methods for selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions are well described in the art. (See, e.g., Ausubel, 1995, supra, ch. 11.)

Typically, oligopeptides 15 residues in length are synthesized using an ABI 431A peptide synthesizer (Perkin-Elmer) using fmoc-chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase immunogenicity. (See, e.g., Ausubel, 1995, supra.) Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for antipeptide activity by, for example, binding the peptide to plastic, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

XIII. Purification of Naturally Occurring HJNCT Using Specific Antibodies

Naturally occurring or recombinant HJNCT is substantially purified by immunoaffinity chromatography using antibodies specific for HJNCT. An immunoaffinity column is constructed by covalently coupling anti-HJNCT antibody to an activated chromatographic resin, such as CNBr-activated SEPHAROSE (Amersham Pharmacia Biotech). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

Media containing HJNCT are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of HJNCT (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/HJNCT binding (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as urea or thiocyanate ion), and HJNCT is collected.

XIV. Identification of Molecules Which Interact with HJNCT

HJNCT, or biologically active fragments thereof, are labeled with ^{125}I Bolton-Hunter reagent. (See, e.g., Bolton, A.E. and W.M. Hunter (1973) Biochem. J. 133:529-539.) Candidate molecules

previously arrayed in the wells of a multi-well plate are incubated with the labeled HJNCT, washed, and any wells with labeled HJNCT complex are assayed. Data obtained using different concentrations of HJNCT are used to calculate values for the number, affinity, and association of HJNCT with the candidate molecules.

Various modifications and variations of the described methods and systems of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with specific preferred embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

Table 1

Polypeptide SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
1	5	2687924	LUNGNOT23	2687924H1 (LUNGNOT23), 549300H1 (BEPINOT01), 2205793F6 (SPLNFET02), 2457286F6 (ENDANOT01), 2664525F6 (ADRENOT08)
2	6	1320134	BLADNOT04	1320134H1 (BLADNOT04), 3373357H1 (CONNTUT05), 881178H1 (THYRNOT02), 2995387H1 (OVARTUT07), 3269474F6 (BRAINOT20), 1320134F1 (BLADNOT04), 4708208H1 (BRAIFET02)
3	7	2594049	OVARTUT02	2594049H1 (OVARTUT02), 2867395T6 (KIDNNOT20), 4071746F6 (KIDNNOT26), SBGA03327F1, SBGA01580F1, SBGA04257F1, SBGA02907F1, SBGA02912F1, SBGA05085F1
4	8	5139028	OVARDIT04	1333953T1 (COLNNOT13), 1843743R6 (COLNNOT08), 1855824F6 (PROSNOT18), 2614186F6 (GBLANOT01)

Table 2

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs, and Domains	Homologous Sequences	Analytical Methods
1	361	T3 T166 S179 S317 S11 T23 S74 T273	N61 N67	GLGF segment (PDZ domain): G90-F93 (+R84) Homer domain: T65-F93 Potential myristoylation sites: G54, G90, G123, G250, G252	Homer [Rattus norvegicus] g1913909	BLAST MOTIFS
2	300	S4 T11 S28 S32 T97 S123 S124 S138 S174 S189 S201 S211 S222 S224 S229 T247 S257 T283 T285 S297	N180	GRASP65 domain: L50-I62 Potential myristoylation sites: G45, G48, G51, G53, G107, G216, G280	GRASP65 [Rattus norvegicus] g4432587	BLAST MOTIFS
3	230	T62 S29 S155 S187 S208 Y224	N190	Transmembrane domains: A2-S29, V117-W138, L164-C182 Epithelial membrane protein domain: F46-T59	Claudin-2 [Mus musculus] g3335184	BLAST BLOCKS HMM MOTIFS
4	292	S176 S7 S59 S288 S24 S165		PDZ domains: E108-R187, T192-S266	Syntenin [Mus musculus] g3342560	BLAST MOTIFS PFAM

Table 3

Nucleotide SEQ ID NO:	Useful Fragments	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
5	572-625	Nervous Reproductive Fetal	Immortalized/Cancer (0.43) Trauma (0.14) Immune response (0.11)	pINCY
6	338-370	Nervous (0.35) Reproductive (0.35)	Cancer/Cell proliferation (0.60) Inflammation/Immune response (0.30) Neurological disorders (0.15)	pINCY
7	272-316	Gastrointestinal (0.211) Urologic (0.211) Reproductive (0.211)	Cell proliferation (0.684) Inflammation (0.263)	pINCY
8	625-669	Gastrointestinal (0.521) Nervous (0.146) Reproductive (0.146)	Cell proliferation (0.563) Inflammation (0.375)	pINCY

Table 4

Nucleotide SEQ ID NO:	Library	Library Comment
5	LUNGNOT23	Library was constructed using RNA isolated from left lobe lung tissue removed from a 58-year-old Caucasian male during a segmental lung resection. Pathology for the associated tumor tissue indicated metastatic grade 3 (of 4) osteosarcoma. Patient history included cancer of the soft tissue, secondary cancer of the lung, prostate cancer, acute duodenal ulcer with hemorrhage, benign hypertension, and radiation therapy. Family history included prostate cancer, benign hypertension, breast cancer, and acute leukemia.
6	BLADNOT04	Library was constructed from bladder tissue obtained from a 28-year-old Caucasian male. The bladder tissue was removed upon death of the individual by a self-inflicted gun shot wound. This library contains markers for seminal vesicle tissue, indicating that it was derived from a heterogeneous mixture of both bladder and seminal vesicle tissue.
7	OVARTUT02	Library was constructed using RNA isolated from ovarian tumor tissue removed from a 51-year-old Caucasian female during an exploratory laparotomy, total abdominal hysterectomy, salpingo-oophorectomy, and an incidental appendectomy. Pathology indicated mucinous cystadenoma presenting as a multiloculated neoplasm involving the entire left ovary. The right ovary contained a follicular cyst and a hemorrhagic corpus luteum. The uterus showed proliferative endometrium and a single intramural leiomyoma. The peritoneal biopsy indicated benign glandular inclusions consistent with endosalpingiosis. Family history included atherosclerotic coronary artery disease, benign hypertension, breast cancer, and uterine cancer.
8	OVARDIT04	Library was constructed using RNA isolated from diseased left ovary tissue removed from a 22-year-old Caucasian female during a left ovarian cystectomy. Pathology indicated a mature cystic teratoma (dermoid cyst) of the left ovary which consisted of aggregate pieces of fibrous tissue and hair. Patient history included capsular disruption of spleen, mononucleosis, and chlamydia (treated at age 21).

Table 5

Program	Description	Reference	Parameter Threshold
ABI FACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	Mismatch <50%
ABI AutoAssembler	A program that assembles nucleic acid sequences.	Perkin-Elmer Applied Biosystems, Foster City, CA.	
BLAST	A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx.	Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. et al. (1997) Nucleic Acids Res. 25: 3389-3402.	*ESTs:* Probability value= 1.0E-8 or less *Full Length sequences:* Probability value= 1.0E-10 or less
FASTA	A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA comprises as least five functions: fasta, tfasta, fastx, tfastx, and ssearch.	Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad Sci. 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183: 63-98; and Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489.	*ESTs:* fasta E value=1.06E-6 *Assembled ESTs:* fasta Identity= 95% or greater and Match length=200 bases or greater; fastx E value=1.0E-8 or less *Full Length sequences:* fastx score=100 or greater
BLIMPS	A BLocks IMProved Searcher that matches a sequence against those in BLOCKS and PRINTS databases to search for gene families, sequence homology, and structural fingerprint regions.	Henikoff, S and J.G. Henikoff, Nucl. Acid Res., 19:6565-72, 1991. J.G. Henikoff and S. Henikoff (1996) Methods Enzymol. 266:88-105; and Attwood, T.K. et al. (1997) J. Chem. Inf. Comput. Sci. 37: 417-424.	Score=1000 or greater; Ratio of Score/Strength = 0.75 or larger; and Probability value= 1.0E-3 or less if applicable
PFAM	A Hidden Markov Models-based application useful for protein family search.	Krogh, A. et al. (1994) J. Mol. Biol., 235:1501-1531; Sonnhammer, E.L.L. et al. (1988) Nucleic Acids Res. 26:320-322.	Score=10-50 bits, depending on individual protein families

Table 5 (cont.)

Program	Description	Reference	Parameter Threshold
ProfileScan	An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite.	Gribskov, M. et al. (1988) CABIOS 4:61-66; Gribskov, et al. (1989) Methods Enzymol. 183:146-159; Bairoch, A. et al. (1997) Nucleic Acids Res. 25: 217-221.	Score= 4.0 or greater
Phred	A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability.	Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	A Phils Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M. S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M. S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	Score= 120 or greater; Match length= 56 or greater
Consed	A graphical tool for viewing and editing Phrap assemblies	Gordon, D. et al. (1998) Genome Res. 8:195-202.	
SPScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nielson, H. et al. (1997) Protein Engineering 10:1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12: 431-439.	Score=5 or greater
Motifs	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Bairoch et al. supra; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

What is claimed is:

1. A substantially purified polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-4 and fragments thereof.

2. A substantially purified variant having at least 95% amino acid sequence identity to the amino acid sequence of claim 1.

3. An isolated and purified polynucleotide encoding the polypeptide of claim 1.

4. An isolated and purified polynucleotide variant having at least 90% polynucleotide sequence identity to the polynucleotide of claim 3.

5. An isolated and purified polynucleotide which hybridizes under stringent conditions to the polynucleotide of claim 3.

6. An isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide of claim 3.

7. A method for detecting a polynucleotide, the method comprising the steps of:

(a) hybridizing the polynucleotide of claim 6 to at least one nucleic acid in a sample, thereby forming a hybridization complex; and

(b) detecting the hybridization complex, wherein the presence of the hybridization complex correlates with the presence of the polynucleotide in the sample.

8. The method of claim 7 further comprising amplifying the polynucleotide prior to hybridization.

9. An isolated and purified polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:5-8 and fragments thereof.

10. An isolated and purified polynucleotide variant having at least 90% polynucleotide sequence identity to the polynucleotide of claim 9.

11. An isolated and purified polynucleotide having a sequence which is complementary to the polynucleotide of claim 9.

12. An expression vector comprising at least a fragment of the polynucleotide of claim 3.

13. A host cell comprising the expression vector of claim 12.

14. A method for producing a polypeptide, the method comprising the steps of:

a) culturing the host cell of claim 13 under conditions suitable for the expression of the polypeptide; and

b) recovering the polypeptide from the host cell culture.

15. A pharmaceutical composition comprising the polypeptide of claim 1 in conjunction with a suitable pharmaceutical carrier.

16. A purified antibody which specifically binds to the polypeptide of claim 1.

17. A purified agonist of the polypeptide of claim 1.

18. A purified antagonist of the polypeptide of claim 1.

19. A method for treating or preventing a disorder associated with decreased expression or activity of HJNCT, the method comprising administering to a subject in need of such treatment an effective amount of the pharmaceutical composition of claim 15.

20. A method for treating or preventing a disorder associated with increased expression or activity of HJNCT, the method comprising administering to a subject in need of such treatment an effective amount of the antagonist of claim 18.

```
1    M S T A R E Q P I F S T R A H V F Q I D  2687924
1    M G K M G E Q P I F S T R A H V F Q I D  g1913909

21   P A T K R N W I P A G K H A L T V S Y F  2687924
21   P N T K K N W V P T S K H A V T V S Y F  g1913909

41   Y D A T R N V Y R I I S I G G A K A I I  2687924
41   Y D S T R N V Y R I I S L D G S K A I I  g1913909

61   N S T V T P N M T F T K T S Q K F G Q W  2687924
61   N S T I T P N M T F T K T S Q K F G Q W  g1913909

81   A D S R A N T V Y G L G F A S E Q H L T  2687924
81   A D S R A N T V Y G L G F S S E H H L S  g1913909

101  Q F A E K F Q E V K E A A R L A R E K S  2687924
101  K F A E K F Q E F K E A A R L A K E K S  g1913909

121  Q D G G E L T S P A L G L A S H Q V P P  2687924
121  Q E K M E L T S T - - - - - - - - - - -  g1913909
```

FIGURE 1A

```
141 S P L V S A N G P G E E K L F R S Q S A 2687924
130 - - - - - - - - - P S Q E - - - - - - - g1913909

161 D A P G P T E R E R L K K M L S E G S V 2687924
134 - - - - - - - - - - - - - - - - - S A G g1913909

181 G E V Q W E A E F F A L Q D S N N K L A 2687924
137 G D L Q - - - - - - - - - - - - - - - - g1913909

201 G A L R E A N A A A A Q W R Q Q L E A Q 2687924
141 - - - - - - - - - - - - - - - - - - - - g1913909

221 R A E A E R L R Q R V A E L E A Q A A S 2687924
141 - - - - - - - - - - - - - - - - - - - S g1913909

241 E V T P T G E K E G L G Q G Q S L E Q L 2687924
142 P L T P - - - - - - - - - - - - - - - - g1913909

261 E A L V Q T K D Q E I Q T L K S Q T G G 2687924
146 - - - - - - - - - - - - - - - - - - - - g1913909
```

FIGURE 1B

```
281 P R E A L E A A E R E E T Q Q K V Q D L 2687924
146 - - E S I N G T D D E R T P D V T Q N S g1913909

301 E T R N A E L E H Q L R A M E R S L E E 2687924
164 E P R A E P A Q N A L P F S H R - - - - g1913909

321 A R A E R E R A R A E V G R A A Q L L D 2687924
180 - - - - - - - - - - - - - - - - - - - - g1913909

341 V R L F E L S E L R E G L A R L A E A A 2687924
180 - - - Y T F N S - - - - - - - - - A I M I g1913909

361 P                                       2687924
189 K                                       g1913909
```

FIGURE 1C

```
1   M A S L G L Q L V G Y I L G L L G L L G T L V A M L L P S W  2594049
1   M A S L G V Q L V G Y I L G L L G L L G T S I A M L L P N W  g3335184

31  K T S S Y V G A S I V T A V G F S K G L W M E C A T H S T G  2594049
31  R T S S Y V G A S I V T A V G F S K G L W M E C A T H S T G  g3335184

61  I T Q C D I Y S T L L G L P A D I Q A A Q A M M V T S S A I  2594049
61  I T Q C D I Y S T L L G L P A D I Q A A Q A M M V T S S A M  g3335184

91  S S L A C I I S V V G M R C T V F C Q E S R A K D R V A V A  2594049
91  S S L A C I I S V V G M R C T V F C Q D S R A K D R V A V V  g3335184

121 G G V F F I L G G L L G F I P V A W N L H G I L R D F Y S P  2594049
121 G G V F F I L G G I L G F I P V A W N L H G I L R D F Y S P  g3335184

151 L V P D S M K F E I G E A L Y L G I I S S L F S L I A G I I  2594049
151 L V P D S M K F E I G E A L Y L G I I S A L F S L V A G V I  g3335184

181 L C F S C S S Q R N R S N Y Y D A Y Q A Q P L A T R S S P R  2594049
181 L C F S C S P Q G N R T N Y Y D G Y Q A Q P L A T R S S P R  g3335184

211 P G Q P P K V K S E F N S Y S L T G Y V                      2594049
211 S A Q Q P K A K S E F N S Y S L T G Y V                      g3335184
```

FIGURE 2

```
1    M S S L Y P S L E D L K V D Q A I Q A Q V R A S P K M P A -  5139028
1    M S - L Y P S L E D L K V D K V I Q A Q T A Y S A N P A S Q  g3342560

30   - - - L P V Q A T A I S P P P V L Y P N L - A E L E N Y M G  5139028
30   A F V L V D A S A A L P P D G N L Y P K L Y P E L S Q Y M G  g3342560

56   L S L S S Q E V Q E S L L Q I P E G D S T A V S G P G P G Q  5139028
60   L S L N E A E I C E S M P M V S G A P A Q G L V A R P S S V  g3342560

86   - - M V A P V T G Y S L G V R R A E I K P G V R E I H L C K  5139028
90   N Y M V A P V T G N D A G I R R A E I K Q G I R E V I L C K  g3342560

114  D E R G K T G L R L R K V D Q G L F V Q L V Q A N T P A S L  5139028
120  D Q D G K I G L R L K S I D N G I F V Q L V Q A N S P A S L  g3342560

144  V G L R F G D Q L L Q I D G R D C A G W S S H K A H Q V V K  5139028
150  V G L R F G D Q V L Q I N G E N C A G W S S D K A H K V L K  g3342560

174  K A S G D K I V M V V R D R P F Q R T V T M H K D S M G H V  5139028
180  Q A F G E K I T M T I R D R P F E R T V I M H K D S S G H V  g3342560

204  G F V I K K G K I V S L V K G S S A A C N G L L T N H Y V C  5139028
210  G F I F K S G K I T S I V K D S S A A R N G L L T D H H I C  g3342560
```

FIGURE 3A

```
234 EVDGQNVIGLKDKKIMEILATAGNVVTLTI 5139028
240 EINGQNVIGLKDAQIADILSTAGTVVTITI g3342560

264 IPSVIYEHMVKKLPPVLLHHTMDHSIPDA  5139028
270 MPTFIFEHIIKRMAPSIMKSLMDHTIPEV  g3342560
```

FIGURE 3B

SEQUENCE LISTING

```
<110> INCYTE PHARMACEUTICALS, INC.
      YUE, Henry
      LAL, Preeti
      CORLEY, Neil C.
      GUEGLER, Karl J.
      BAUGHN, Mariah R.
      LU, Aina D.
      TANG, Y. Tom

<120> MEMBRANE-ASSOCIATED ORGANIZATIONAL PROTEINS

<130> PF-0590 PCT

<140> To Be Assigned
<141> Herewith

<150> 09/161,043; unassigned; 09/172,189; unassigned; 09/305,329;
      unassigned
<151> 1998-09-25; 1998-09-25; 1999-10-13; 1999-10-13; 1999-05-04;
      1999-05-04

<160> 12

<170> PERL Program

<210> 1
<211> 361
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2687924CD1

<400> 1
Met Ser Thr Ala Arg Glu Gln Pro Ile Phe Ser Thr Arg Ala His
  1               5                  10                  15
Val Phe Gln Ile Asp Pro Ala Thr Lys Arg Asn Trp Ile Pro Ala
                 20                  25                  30
Gly Lys His Ala Leu Thr Val Ser Tyr Phe Tyr Asp Ala Thr Arg
                 35                  40                  45
Asn Val Tyr Arg Ile Ile Ser Ile Gly Gly Ala Lys Ala Ile Ile
                 50                  55                  60
Asn Ser Thr Val Thr Pro Asn Met Thr Phe Thr Lys Thr Ser Gln
                 65                  70                  75
Lys Phe Gly Gln Trp Ala Asp Ser Arg Ala Asn Thr Val Tyr Gly
                 80                  85                  90
Leu Gly Phe Ala Ser Glu Gln His Leu Thr Gln Phe Ala Glu Lys
                 95                 100                 105
Phe Gln Glu Val Lys Glu Ala Ala Arg Leu Ala Arg Glu Lys Ser
                110                 115                 120
Gln Asp Gly Gly Glu Leu Thr Ser Pro Ala Leu Gly Leu Ala Ser
```

```
                125                     130                     135
His Gln Val Pro Pro Ser Pro Leu Val Ser Ala Asn Gly Pro Gly
                140                     145                     150
Glu Glu Lys Leu Phe Arg Ser Gln Ser Ala Asp Ala Pro Gly Pro
                155                     160                     165
Thr Glu Arg Glu Arg Leu Lys Lys Met Leu Ser Glu Gly Ser Val
                170                     175                     180
Gly Glu Val Gln Trp Glu Ala Glu Phe Phe Ala Leu Gln Asp Ser
                185                     190                     195
Asn Asn Lys Leu Ala Gly Ala Leu Arg Glu Ala Asn Ala Ala Ala
                200                     205                     210
Ala Gln Trp Arg Gln Gln Leu Glu Ala Gln Arg Ala Glu Ala Glu
                215                     220                     225
Arg Leu Arg Gln Arg Val Ala Glu Leu Glu Ala Gln Ala Ala Ser
                230                     235                     240
Glu Val Thr Pro Thr Gly Glu Lys Glu Gly Leu Gly Gln Gly Gln
                245                     250                     255
Ser Leu Glu Gln Leu Glu Ala Leu Val Gln Thr Lys Asp Gln Glu
                260                     265                     270
Ile Gln Thr Leu Lys Ser Gln Thr Gly Gly Pro Arg Glu Ala Leu
                275                     280                     285
Glu Ala Ala Glu Arg Glu Glu Thr Gln Gln Lys Val Gln Asp Leu
                290                     295                     300
Glu Thr Arg Asn Ala Glu Leu Glu His Gln Leu Arg Ala Met Glu
                305                     310                     315
Arg Ser Leu Glu Glu Ala Arg Ala Glu Arg Glu Arg Ala Arg Ala
                320                     325                     330
Glu Val Gly Arg Ala Ala Gln Leu Leu Asp Val Arg Leu Phe Glu
                335                     340                     345
Leu Ser Glu Leu Arg Glu Gly Leu Ala Arg Leu Ala Glu Ala Ala
                350                     355                     360
Pro
```

<210> 2
<211> 300
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1320134CD1

<400> 2

```
Met Arg Phe Ser Gln Ser Glu Asp Phe Phe Thr Leu Ile Glu Ser
  1             5                    10                      15
His Glu Gly Lys Pro Leu Lys Leu Met Val Tyr Asn Ser Lys Ser
                20                      25                      30
Asp Ser Cys Arg Glu Val Thr Val Thr Pro Asn Ala Ala Trp Gly
                35                      40                      45
Gly Glu Gly Ser Leu Gly Cys Gly Ile Gly Tyr Gly Tyr Leu His
                50                      55                      60
Arg Ile Pro Thr Gln Pro Pro Ser Tyr His Lys Lys Pro Pro Gly
                65                      70                      75
Thr Pro Pro Pro Ser Ala Leu Pro Leu Gly Ala Pro Pro Pro Asp
```

```
                    80                  85                  90
Ala Leu Pro Pro Gly Pro Thr Pro Glu Asp Ser Pro Ser Leu Glu
                    95                  100                 105
Thr Gly Ser Arg Gln Ser Asp Tyr Met Glu Ala Leu Leu Gln Ala
                    110                 115                 120
Pro Gly Ser Ser Met Glu Asp Pro Leu Pro Gly Pro Gly Ser Pro
                    125                 130                 135
Ser His Ser Ala Pro Asp Pro Asp Gly Leu Pro His Phe Met Glu
                    140                 145                 150
Thr Pro Leu Gln Pro Pro Pro Pro Val Gln Arg Val Met Asp Pro
                    155                 160                 165
Gly Phe Leu Asp Val Ser Gly Ile Ser Leu Leu Asp Asn Ser Asn
                    170                 175                 180
Ala Ser Val Trp Pro Ser Leu Pro Ser Ser Thr Glu Leu Thr Thr
                    185                 190                 195
Thr Ala Val Ser Thr Ser Gly Pro Glu Asp Ile Cys Ser Ser Ser
                    200                 205                 210
Ser Ser His Glu Arg Gly Gly Glu Ala Thr Trp Ser Gly Ser Glu
                    215                 220                 225
Phe Glu Val Ser Phe Leu Asp Ser Pro Gly Ala Gln Ala Gln Ala
                    230                 235                 240
Asp His Leu Pro Gln Leu Thr Leu Pro Asp Ser Leu Thr Ser Ala
                    245                 250                 255
Ala Ser Pro Glu Asp Gly Leu Ser Ala Glu Leu Leu Glu Ala Gln
                    260                 265                 270
Ala Glu Glu Glu Pro Ala Ser Thr Glu Gly Leu Asp Thr Gly Thr
                    275                 280                 285
Glu Ala Glu Gly Leu Asp Ser Gln Ala Gln Ile Ser Thr Thr Glu
                    290                 295                 300


<210> 3
<211> 230
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2594049CD1

<400> 3
Met Ala Ser Leu Gly Leu Gln Leu Val Gly Tyr Ile Leu Gly Leu
  1                 5                   10                  15
Leu Gly Leu Leu Gly Thr Leu Val Ala Met Leu Leu Pro Ser Trp
                    20                  25                  30
Lys Thr Ser Ser Tyr Val Gly Ala Ser Ile Val Thr Ala Val Gly
                    35                  40                  45
Phe Ser Lys Gly Leu Trp Met Glu Cys Ala Thr His Ser Thr Gly
                    50                  55                  60
Ile Thr Gln Cys Asp Ile Tyr Ser Thr Leu Leu Gly Leu Pro Ala
                    65                  70                  75
Asp Ile Gln Ala Ala Gln Ala Met Met Val Thr Ser Ser Ala Ile
                    80                  85                  90
Ser Ser Leu Ala Cys Ile Ile Ser Val Val Gly Met Arg Cys Thr
                    95                  100                 105
```

```
Val Phe Cys Gln Glu Ser Arg Ala Lys Asp Arg Val Ala Val Ala
                110                 115                 120
Gly Gly Val Phe Phe Ile Leu Gly Gly Leu Leu Gly Phe Ile Pro
                125                 130                 135
Val Ala Trp Asn Leu His Gly Ile Leu Arg Asp Phe Tyr Ser Pro
                140                 145                 150
Leu Val Pro Asp Ser Met Lys Phe Glu Ile Gly Glu Ala Leu Tyr
                155                 160                 165
Leu Gly Ile Ile Ser Ser Leu Phe Ser Leu Ile Ala Gly Ile Ile
                170                 175                 180
Leu Cys Phe Ser Cys Ser Ser Gln Arg Asn Arg Ser Asn Tyr Tyr
                185                 190                 195
Asp Ala Tyr Gln Ala Gln Pro Leu Ala Thr Arg Ser Ser Pro Arg
                200                 205                 210
Pro Gly Gln Pro Pro Lys Val Lys Ser Glu Phe Asn Ser Tyr Ser
                215                 220                 225
Leu Thr Gly Tyr Val
                230


<210> 4
<211> 292
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5139028CD1

<400> 4
Met Ser Ser Leu Tyr Pro Ser Leu Glu Asp Leu Lys Val Asp Gln
  1               5                  10                  15
Ala Ile Gln Ala Gln Val Arg Ala Ser Pro Lys Met Pro Ala Leu
                 20                  25                  30
Pro Val Gln Ala Thr Ala Ile Ser Pro Pro Pro Val Leu Tyr Pro
                 35                  40                  45
Asn Leu Ala Glu Leu Glu Asn Tyr Met Gly Leu Ser Leu Ser Ser
                 50                  55                  60
Gln Glu Val Gln Glu Ser Leu Leu Gln Ile Pro Glu Gly Asp Ser
                 65                  70                  75
Thr Ala Val Ser Gly Pro Gly Pro Gly Gln Met Val Ala Pro Val
                 80                  85                  90
Thr Gly Tyr Ser Leu Gly Val Arg Arg Ala Glu Ile Lys Pro Gly
                 95                 100                 105
Val Arg Glu Ile His Leu Cys Lys Asp Glu Arg Gly Lys Thr Gly
                110                 115                 120
Leu Arg Leu Arg Lys Val Asp Gln Gly Leu Phe Val Gln Leu Val
                125                 130                 135
Gln Ala Asn Thr Pro Ala Ser Leu Val Gly Leu Arg Phe Gly Asp
                140                 145                 150
Gln Leu Leu Gln Ile Asp Gly Arg Asp Cys Ala Gly Trp Ser Ser
                155                 160                 165
His Lys Ala His Gln Val Val Lys Lys Ala Ser Gly Asp Lys Ile
                170                 175                 180
Val Met Val Val Arg Asp Arg Pro Phe Gln Arg Thr Val Thr Met
```

```
                185                 190                 195
His Lys Asp Ser Met Gly His Val Gly Phe Val Ile Lys Lys Gly
                200                 205                 210
Lys Ile Val Ser Leu Val Lys Gly Ser Ser Ala Ala Cys Asn Gly
                215                 220                 225
Leu Leu Thr Asn His Tyr Val Cys Glu Val Asp Gly Gln Asn Val
                230                 235                 240
Ile Gly Leu Lys Asp Lys Lys Ile Met Glu Ile Leu Ala Thr Ala
                245                 250                 255
Gly Asn Val Val Thr Leu Thr Ile Ile Pro Ser Val Ile Tyr Glu
                260                 265                 270
His Met Val Lys Lys Leu Pro Pro Val Leu Leu His His Thr Met
                275                 280                 285
Asp His Ser Ile Pro Asp Ala
                290
```

<210> 5
<211> 1501
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2687924CB1

<400> 5

```
gcacgaaggg agtccccgga gttaccgacc tggggttgga ggaggcagaa gaggtgtccc   60
ccgaagttct ggacccagag acagggggcg tcgagggaat ctcggaagtt ccagggggcc  120
aaaccagtgc tcctgccacc tctctggctg ccccctagag cctgcccatc ccagcctgac  180
caatgtccac agccagggag cagccaatct tcagcacacg ggcgcacgtg ttccaaattg  240
acccagccac caagcgaaac tggatcccag cgggcaagca cgcactcact gtctcctatt  300
tctacgatgc cacccgcaat gtgtaccgca tcatcagcat cggaggcgcc aaggccatca  360
tcaacagcac tgtcactccc aacatgacct tcaccaaaac ttcccagaag ttcgggcagt  420
gggccgacag tcgcgccaac acagtctacg gcctgggctt tgcctctgaa cagcatctga  480
cacagtttgc cgagaagttc caggaagtga aggaagcagc caggctggcc agggagaaat  540
ctcaggatgg cggggagctc accagtccag ccctggggct cgcctcccac caggtgcccc  600
cgagccctct cgtcagtgcc aacggccccg gcgaggaaaa actgttccgc agccagagcg  660
ctgatgcccc cggcccccaca gagcgcgagc ggctaaagaa gatgttgtct gagggctccg  720
tgggcgaggt acagtgggag gccgagtttt tcgcactgca ggacagcaac aacaagctgg  780
caggcgccct gcgagaggcc aacgccgccg cagcccagtg gaggcagcag ctggaggctc  840
agcgtgcaga ggccgagcgg ctgcggcagc gggtggctga gctggaggct caggcagctt  900
cagaggtgac ccccaccggt gagaaggagg ggctgggcca gggccagtcg ctggaacagc  960
tggaagctct ggtgcaaacc aaggaccagg agattcagac cctgaagagt cagactgggg 1020
ggccccgcga ggccctggag gctgccgagc gtgaggagac tcagcagaag gtgcaggacc 1080
tggagacccg caatgcggag ttggagcacc agctgcgggc gatggagcgc agcctggagg 1140
aggcacgggc agagcgggag cgggcgcggg ctgaggtggg ccgggcagcg cagctgctgg 1200
acgtcaggct gtttgagctg agtgagctgc gtgagggcct ggcccgcctg gctgaggctg 1260
cgccctgagc cggggctggt tttctatgaa cgattccggc ctgggatgcg ggccaggctg 1320
caggcggcat agttgggccc attcgtcctg gaaagggact gggggtcccc aacttagccc 1380
tgggtgggcc gggccgggct gggctggggt gggccccagt cggctctggt tgttggcagc 1440
tttggggctg tttttgagct tctcattgtg tagaatttct agatccccccg attacatttc 1500
g                                                                 1501
```

<210> 6
<211> 1039
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1320134CB1

<400> 6
```
ctggatgtgg tgtgaagaga agggtgctgt ccatgagaag tgtctggttt tgttgaggct   60
ttagccctgg aagcaggcct caggtgctat gcgtttctca cagtccgagg acttctttac  120
gctcatcgag tctcatgagg ggaagccctt gaagctgatg gtgtataact ccaagtcaga  180
ctcctgccgg gaggtgactg taactcccaa cgcagcctgg ggtggagagg gcagtctggg  240
atgtggcatt ggctatgggt atctacaccg gatcccaact cagcccccca gctaccacaa  300
gaagccacct ggcaccccac caccttctgc tctaccactt ggtgccccac cacctgatgc  360
tctaccacct ggacccaccc ccgaggactc tccttccctg gagacaggtt ccaggcagag  420
tgactacatg gaggccctgc tgcaggcacc tggctcctcc atggaggatc cccttcctgg  480
gcctgggagt cccagccaca gtgctccaga ccctgatgga cttccccatt tcatggagac  540
tcctcttcag cccccacctc cagtgcagcg agttatggac ccaggcttcc tggacgtgtc  600
gggaatttct ctcttggaca acagcaatgc cagtgtgtgg cccagcctgc cctcttccac  660
agaactgacc accacagctg tctcaacctc agggccagag gacatctgct ccagcagcag  720
ttctcatgag cggggtggtg aggctacatg gtctgggtca gagtttgagg tctccttcct  780
ggacagccca ggtgcccaag cccaggcgga ccacctgcct cagctgactc ttcctgacag  840
tctcacctct gcagcctcac cagaagatgg gctgtccgcc gagctgcttg aagctcaggc  900
tgaggaggaa ccagcaagca cagagggcct agatactggg acggaggctg aggggctgga  960
cagccaggcc cagatctcta ccacagaata acaccctggg ctgtgacaag gcccatgatg 1020
acatttcatg aggcccaga                                              1039
```

<210> 7
<211> 2742
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2594049CB1

<400> 7
```
agaagtcagc ctggcagaga gactctgaaa tgagggatta gaggtgttca aggagcaaga   60
gcttcagcct gaagacaagg gagcagtccc tgaagacgct tctactgaga ggtctgccat  120
ggcctctctt ggcctccaac ttgtgggcta catcctaggc cttctggggc ttttgggcac  180
actggttgcc atgctgctcc ccagctggaa aacaagttct tatgtcggtg ccagcattgt  240
gacagcagtt ggcttctcca agggcctctg gatggaatgt gccacacaca gcacaggcat  300
cacccagtgt gacatctata gcacccttct gggcctgccc gctgacatcc aggctgccca  360
ggccatgatg gtgacatcca gtgcaatctc ctccctggcc tgcattatct ctgtggtggg  420
catgagatgc acagtcttct gccaggaatc ccgagccaaa gacagagtgg cggtagcagg  480
tggagtcttt ttcatccttg gaggcctcct gggattcatt cctgttgcct ggaatcttca  540
tgggatccta cgggacttct actcaccact ggtgcctgac agcatgaaat ttgagattgg  600
agaggctctt tacttgggca ttatttcttc cctgttctcc ctgatagctg gaatcatcct  660
ctgcttttcc tgctcatccc agagaaatcg ctccaactac tacgatgcct accaagccca  720
acctcttgcc acaaggagct ctccaaggcc tggtcaacct cccaaagtca agagtgagtt  780
caattcctac agcctgacag ggtatgtgtg aagaaccagg ggccagagct gggggggtggc 840
tgggtctgtg aaaaacagtg gacagcaccc cgagggccac aggtgaggga cactaccact  900
```

```
ggatcgtgtc agaaggtgct gctgaggata gactgacttt ggccattgga ttgagcaaag 960
gcagaaatgg gggctagtgt aacagcatgc aggttgaatt gccaaggatg ctcgccatgc 1020
cagcctttct gttttcctca ccttgctgct cccctgccct aagtccccaa ccctcaactt 1080
gaaaccccat tcccttaagc caggactcag aggatccctt tgccctctgg tttacctggg 1140
actccatccc caaacccact aatcacatcc cactgactga ccctctgtga tcaaagaccc 1200
tctctctggc tgaggttggc tcttagctca ttgctgggga tgggaaggag aagcagtggc 1260
ttttgtgggc attgctctaa cctacttctc aagcttccct ccaaagaaac tgattggccc 1320
tggaacctcc atcccactct tgttatgact ccacagtgtc cagactaatt tgtgcatgaa 1380
ctgaaataaa accatcctac ggtatccagg gaacagaaag caggatgcag gatgggagga 1440
caggaaggca gcctgggaca tttaaaaaaa taaaaatgaa aaaaaaaccc agaacccatt 1500
tctcagggca ctttccagaa ttctctcata tttgtgggct gggatcaagc ctgcagcttg 1560
aggaaagcac aaggaaagga aagaagatct ggtggaaagc tcaggtggca gcggactctg 1620
actccactga ggaactgcct cagaagctgc gatcacaact ttggctgaag cccctgcctc 1680
actctagggc acctgacctg gcctcttgcc taaaccacaa ggctaagggc tatagacaat 1740
ggtttcctta ggaacagtaa accagttttt ctagggatgg cccttggctg gggggatgaca 1800
gtgtgggagc tgtggggtac tgaggaagac accattcctt gacggtgtct aagaagccag 1860
gtggatgtgt gtggtggctc cagtgggtgt ttctactctg ccagtgagag gcagcccccct 1920
agaaactctt caggcgtaat ggaaaatcag ctcaaatgag atcaggcccc cccagggtcc 1980
acccacagag cactacagag cctctgaaag accatagcac caagcgagcc ccttcagatt 2040
cccccactgt ccatcggaag atgctccaga gtggctagag ggcatctaag ggctccagca 2100
tggcatatcc atgcccacgg tgctgtgtcc atgatctgag tgatagctgc actgctgcct 2160
gggattgcag ctgaggtggg agtggagaat ggttcccagg aagacagttc cacctctaag 2220
gtccgaaaat gttcccttta ccctggagtg ggagtgaggg gtcatacacc aaaggtattt 2280
tccctcacca gtctaggcat gactggcttc tgaaaaattc cagcacacct cctcgaacct 2340
cattgtcagc agagagggcc catctgttgt ctgtaacatg cctttcacat gtccaccttc 2400
ttgccatgtt ccagctgctc tcccaacctg gaaggccgtc tcccccttagc caagtcctcc 2460
tcaggcttgg agaacttcct cagcgtcacc tccttcattg agccttctct gatcactcca 2520
tccctctcct acccctccct cccccaaccc tcaatgtata aattgcttct tgatgcttag 2580
cattcacaat ttttgattga tcgttatttg tgtgtgtgtg tccgatctca caagtatatt 2640
gtaaaccctt cggtgggtgg gggccatatc ctagacctct ctgtatcccc cagactatct 2700
gtaacagtgc caggcaccag aaggtgatca ataattcgta gc                   2742
```

<210> 8
<211> 1473
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5139028CB1

<400> 8

```
ggaagccagc tgagagacag ggctacgttt cagggaggga aacagattca gcagcggcag   60
cagctggaga aggtcgtgga gcagcacctt gcctgcaggg tgttctgaga atcagccatg 120
tcatccctgt acccatctct agaggaccta aaagtggacc aagccattca ggcccaggtc 180
agagcctcac ccaagatgcc agccctgcca gtccaggcaa cagccatttc cccaccacca 240
gttttgtacc caaacttggc agaactggaa aattatatgg gtctttccct ctccagccaa 300
gaagtccagg agagcctgct tcagattcca gagggtgaca gtacagcggt ctcgggcccc 360
gggcccggcc agatggtggc accggtaacc gggtacagcc tgggcgtgcg gcgagctgag 420
atcaagcccg gggtgcgcga gatccacctg tgcaaggacg agcgcggcaa gaccgggctg 480
aggctgcgga aggtcgacca ggggctcttt gtgcagttgg tccaggccaa cacccctgca 540
tcccttgtgg ggctgcgctt tggggaccag ctcctgcaga ttgacgggcg tgactgtgct 600
gggtggagct cgcacaaagc ccatcaggtg gtgaagaagg catcaggcga taagattgtc 660
atggtggttc gggacaggcc gttccagcgg actgtcacca tgcacaagga cagcatgggc 720
```

```
cacgtcggct tcgtgatcaa gaaggggaag attgtctctc tggtcaaagg gagttctgcg 780
gcctgcaacg ggctcctcac caaccactac gtgtgtgagg tggacgggca gaatgttatc 840
gggctgaagg acaaaaagat catggagatt ctggccacgg ctgggaacgt tgtcaccctg 900
accatcatcc ccagtgtgat ctacgagcac atggtcaaaa agttgcctcc agtcctgctc 960
caccacacca tggaccactc catcccagat gcctgaagcc actgcagggc agggcaggca 1020
ggggggcttt cccgccctcc tgcagcaaag ggcaaccacc ctcggatgat gggttgcagc 1080
cggcctgctg cttaaggtgg gggctgccat gagggggcg tgtccaggag ggtgaccatg 1140
ggatggctta tacacacagg cctccttgga gcctcagact ccaagctagg ctgaggctca 1200
ggcagggccc acaggcagcc gattctcttg tgctgattta aatgctggac acggaggcag 1260
gctgtttaaa cgctgcttaa agtcgcaact gggccccttt caagaaattt tgctctacca 1320
ggaaaacagt tacacatttt aagagaacag agctacgttc tttgtgagag ctttttcctt 1380
gccttgactt gctctttgtc acagactgca taagttgtca gccttgacta tcttttgaat 1440
aaagatttga ttttaaacaa aaaaaaaaaa aaa                             1473
```

```
<210> 9
<211> 189
<212> PRT
<213> Rattus norvegicus

<300>
<308> GenBank ID No: g1913909

<400> 9
Met Gly Lys Met Gly Glu Gln Pro Ile Phe Ser Thr Arg Ala His
 1               5                  10                  15
Val Phe Gln Ile Asp Pro Asn Thr Lys Lys Asn Trp Val Pro Thr
                20                  25                  30
Ser Lys His Ala Val Thr Val Ser Tyr Phe Tyr Asp Ser Thr Arg
                35                  40                  45
Asn Val Tyr Arg Ile Ile Ser Leu Asp Gly Ser Lys Ala Ile Ile
                50                  55                  60
Asn Ser Thr Ile Thr Pro Asn Met Thr Phe Thr Lys Thr Ser Gln
                65                  70                  75
Lys Phe Gly Gln Trp Ala Asp Ser Arg Ala Asn Thr Val Tyr Gly
                80                  85                  90
Leu Gly Phe Ser Ser Glu His His Leu Ser Lys Phe Ala Glu Lys
                95                 100                 105
Phe Gln Glu Phe Lys Glu Ala Ala Arg Leu Ala Lys Glu Lys Ser
               110                 115                 120
Gln Glu Lys Met Glu Leu Thr Ser Thr Pro Ser Gln Glu Ser Ala
               125                 130                 135
Gly Gly Asp Leu Gln Ser Pro Leu Thr Pro Glu Ser Ile Asn Gly
               140                 145                 150
Thr Asp Asp Glu Arg Thr Pro Asp Val Thr Gln Asn Ser Glu Pro
               155                 160                 165
Arg Ala Glu Pro Ala Gln Asn Ala Leu Pro Phe Ser His Arg Tyr
               170                 175                 180
Thr Phe Asn Ser Ala Ile Met Ile Lys
               185
```

<210> 10
<211> 451
<212> PRT
<213> Rattus norvegicus

<300>
<308> GenBank ID No: g4432587

<400> 10
Met Gly Leu Gly Ala Ser Ser Glu Gln Pro Ala Gly Gly Glu Gly
1 5 10 15
Phe His Leu His Gly Val Gln Glu Asn Ser Pro Ala Gln Gln Ala
20 25 30
Gly Leu Glu Pro Tyr Phe Asp Phe Ile Ile Thr Ile Gly His Ser
35 40 45
Arg Leu Asn Lys Glu Asn Asp Thr Leu Lys Ala Leu Leu Lys Ala
50 55 60
Asn Val Glu Lys Pro Val Lys Leu Glu Val Phe Asn Met Lys Thr
65 70 75
Met Arg Val Arg Glu Val Glu Val Val Pro Ser Asn Met Trp Gly
80 85 90
Gly Gln Gly Leu Leu Gly Ala Ser Val Arg Phe Cys Ser Phe Arg
95 100 105
Arg Ala Ser Glu His Val Trp His Val Leu Asp Val Glu Pro Ser
110 115 120
Ser Pro Ala Ala Leu Ala Gly Leu Arg Pro Tyr Thr Asp Tyr Ile
125 130 135
Val Gly Ser Asp Gln Ile Leu Gln Glu Ser Glu Asp Phe Phe Thr
140 145 150
Leu Ile Glu Ser His Glu Gly Lys Pro Leu Lys Leu Met Val Tyr
155 160 165
Asn Ser Glu Ser Asp Ser Cys Arg Glu Val Thr Val Thr Pro Asn
170 175 180
Ala Ala Trp Gly Gly Glu Gly Ser Leu Gly Cys Gly Ile Gly Tyr
185 190 195
Gly Tyr Leu His Arg Ile Pro Thr Gln Pro Ser Ser Gln Tyr Lys
200 205 210
Lys Pro Pro Ser Ala Ser Ser Pro Gly Thr Pro Ala Lys Thr Pro
215 220 225
Gln Pro Asn Ala Phe Pro Leu Gly Ala Pro Pro Pro Trp Pro Ile
230 235 240
Pro Gln Asp Ser Ser Gly Pro Glu Leu Gly Ser Arg Gln Ser Asp
245 250 255
Tyr Met Glu Ala Leu Pro Gln Val Pro Gly Gly Phe Met Glu Glu
260 265 270
Gln Leu Pro Gly Pro Gly Ser Pro Gly His Gly Thr Ala Asp Tyr
275 280 285
Gly Gly Cys Leu His Ser Met Glu Ile Pro Leu Gln Pro Pro Pro
290 295 300
Pro Val Gln Arg Val Met Asp Pro Gly Phe Leu Asp Val Ser Gly
305 310 315
Met Ser Leu Leu Asp Ser Asn Asn Thr Ser Val Cys Pro Ser Leu
320 325 330
Ser Ser Ser Ser Leu Leu Thr Pro Thr Ala Val Ser Ala Leu Gly
335 340 345

Pro Glu Asp Ile Gly Ser Ser Ser Ser Ser His Glu Arg Gly Gly
350 355 360
Glu Ala Thr Trp Ser Gly Ser Glu Phe Glu Ile Ser Phe Pro Asp
365 370 375
Ser Pro Gly Ser Gln Ala Gln Val Asp His Leu Pro Arg Leu Thr
380 385 390
Leu Pro Asp Gly Leu Thr Ser Ala Ala Ser Pro Glu Glu Gly Leu
395 400 405
Ser Ala Glu Leu Leu Glu Ala Gln Thr Glu Glu Pro Ala His Thr
410 415 420
Ala Ser Leu Asp Cys Met Ala Gln Thr Glu Gly Pro Ala Gly Gln
425 430 435
Val Gln Ala Ala Pro Asp Pro Glu Pro Gly Leu Cys Glu Gly Pro
440 445 450
Trp

<210> 11
<211> 230
<212> PRT
<213> Mus musculus

<300>
<308> GenBank ID No: g3335184

<400> 11
Met Ala Ser Leu Gly Val Gln Leu Val Gly Tyr Ile Leu Gly Leu
1 5 10 15
Leu Gly Leu Leu Gly Thr Ser Ile Ala Met Leu Leu Pro Asn Trp
20 25 30
Arg Thr Ser Ser Tyr Val Gly Ala Ser Ile Val Thr Ala Val Gly
35 40 45
Phe Ser Lys Gly Leu Trp Met Glu Cys Ala Thr His Ser Thr Gly
50 55 60
Ile Thr Gln Cys Asp Ile Tyr Ser Thr Leu Leu Gly Leu Pro Ala
65 70 75
Asp Ile Gln Ala Ala Gln Ala Met Met Val Thr Ser Ser Ala Met
80 85 90
Ser Ser Leu Ala Cys Ile Ile Ser Val Val Gly Met Arg Cys Thr
95 100 105
Val Phe Cys Gln Asp Ser Arg Ala Lys Asp Arg Val Ala Val Val
110 115 120
Gly Gly Val Phe Phe Ile Leu Gly Gly Ile Leu Gly Phe Ile Pro
125 130 135
Val Ala Trp Asn Leu His Gly Ile Leu Arg Asp Phe Tyr Ser Pro
140 145 150
Leu Val Pro Asp Ser Met Lys Phe Glu Ile Gly Glu Ala Leu Tyr
155 160 165
Leu Gly Ile Ile Ser Ala Leu Phe Ser Leu Val Ala Gly Val Ile
170 175 180
Leu Cys Phe Ser Cys Ser Pro Gln Gly Asn Arg Thr Asn Tyr Tyr
185 190 195
Asp Gly Tyr Gln Ala Gln Pro Leu Ala Thr Arg Ser Ser Pro Arg
200 205 210

Ser Ala Gln Gln Pro Lys Ala Lys Ser Glu Phe Asn Ser Tyr Ser
215 220 225
Leu Thr Gly Tyr Val
230

<210> 12
<211> 298
<212> PRT
<213> Mus musculus

<300>
<308> GenBank ID No: g3342560

<400> 12
Met Ser Leu Tyr Pro Ser Leu Glu Asp Leu Lys Val Asp Lys Val
1 5 10 15
Ile Gln Ala Gln Thr Ala Tyr Ser Ala Asn Pro Ala Ser Gln Ala
20 25 30
Phe Val Leu Val Asp Ala Ser Ala Ala Leu Pro Pro Asp Gly Asn
35 40 45
Leu Tyr Pro Lys Leu Tyr Pro Glu Leu Ser Gln Tyr Met Gly Leu
50 55 60
Ser Leu Asn Glu Ala Glu Ile Cys Glu Ser Met Pro Met Val Ser
65 70 75
Gly Ala Pro Ala Gln Gly Leu Val Ala Arg Pro Ser Ser Val Asn
80 85 90
Tyr Met Val Ala Pro Val Thr Gly Asn Asp Ala Gly Ile Arg Arg
95 100 105
Ala Glu Ile Lys Gln Gly Ile Arg Glu Val Ile Leu Cys Lys Asp
110 115 120
Gln Asp Gly Lys Ile Gly Leu Arg Leu Lys Ser Ile Asp Asn Gly
125 130 135
Ile Phe Val Gln Leu Val Gln Ala Asn Ser Pro Ala Ser Leu Val
140 145 150
Gly Leu Arg Phe Gly Asp Gln Val Leu Gln Ile Asn Gly Glu Asn
155 160 165
Cys Ala Gly Trp Ser Ser Asp Lys Ala His Lys Val Leu Lys Gln
170 175 180
Ala Phe Gly Glu Lys Ile Thr Met Thr Ile Arg Asp Arg Pro Phe
185 190 195
Glu Arg Thr Val Ile Met His Lys Asp Ser Ser Gly His Val Gly
200 205 210
Phe Ile Phe Lys Ser Gly Lys Ile Thr Ser Ile Val Lys Asp Ser
215 220 225
Ser Ala Ala Arg Asn Gly Leu Leu Thr Asp His His Ile Cys Glu
230 235 240
Ile Asn Gly Gln Asn Val Ile Gly Leu Lys Asp Ala Gln Ile Ala
245 250 255
Asp Ile Leu Ser Thr Ala Gly Thr Val Val Thr Ile Thr Ile Met
260 265 270
Pro Thr Phe Ile Phe Glu His Ile Ile Lys Arg Met Ala Pro Ser
275 280 285
Ile Met Lys Ser Leu Met Asp His Thr Ile Pro Glu Val
290 295